Mariângela Spotti Lopes Fujita
Dulce Amélia de Brito Neves
Paula Regina Dal'Evedove
(organizadoras)

Leitura Documentária:

estudos avançados para a indexação

CULTURA ACADÊMICA
Editora

Leitura documentária

Mariângela Spotti Lopes Fujita
Dulce Amélia de Brito Neves
Paula Regina Dal'Evedove
(Organizadoras)

Leitura documentária: estudos avançados para a indexação

Marília/Oficina Universitária
São Paulo/Cultura Acadêmica

2017

Sumário

Apresentação

O documento e suas leituras

Mariângela Spotti Lopes Fujita
Dulce Amélia de Brito Neves
Paula Regina Dal'Evedove

Vivemos leituras diversas em nossa vida cotidiana em casa, no trabalho, no mundo... Essas leituras são interpretações baseadas em nossa história de vida, em nossas crenças e afirmações.

A leitura faz parte da nossa vida como um hábito trivial, cuja implementação é de fácil aplicação. Na perspectiva do senso comum**,** é considerada como a simples junção das letras do alfabeto para a formação de palavras, um ato automático, mecânico (NEVES, 2004, p.34).

Lemos o mundo dentro de uma perspectiva pessoal interpretando-o de acordo com nossa idade, vivência, escolaridade e leituras informacionais do mundo que nos rodeia.

Mesmo a leitura documentária uma técnica aplicada a extração de conceitos importantes e significativos no contexto de uma unidade de

informação (biblioteca, arquivo e/ou museu), não se exime direcionar-se para um "leitor-modelo", aquele que usa o sistema.

Com essa necessidade é que definimos o documentalista (bibliotecário, arquivista, museólogo) como leitor preferencial desse livro, cujo objeto de trabalho é o documento com objetivo de recuperação do conteúdo por um leitor do sistema de sua unidade de informação. Considerando que o documentalista possa estar em formação inicial na graduação ou na pós-graduação, recomendamos a leitura desse livro aos professores e pesquisadores.

Nessa perspectiva, consideramos pertinente mencionar que as pesquisas e publicações sobre leitura documentária em Ciência da Informação têm abordagens teóricas metodológicas cognitivas, semióticas, lingüísticas e psicopedagógicas por ser a leitura um processo mental associado ao pensamento, memória, conhecimento e informação de um indivíduo. Muitas tarefas de processamento e tratamento da informação de documentos são realizadas pela leitura para execução de procedimentos de indexação, classificação, elaboração de resumos que exigem compreensão para decisões, como por exemplo, a identificação de conceitos que representam conteúdos para futura recuperação pelo leitor da unidade de informação.

Esse leitor que realiza a leitura documentária é a principal variável de investigação na abordagem cognitiva e psicopedagógica, porém, a leitura é documentária não somente porque seus fins são a representação e recuperação documentárias, mas também porque o documento é o objeto da leitura e, para isso, as abordagens semiótica e lingüística propiciam conhecimentos sobre o texto e o contexto, as outras variáveis da leitura documentária.

Nas variáveis texto e contexto podemos afirmar que existe grande diversidade de tipos e estruturas e nisso reside, em muitos casos, a complexidade da leitura. Ler um romance é diferente de ler um texto científico quando o contexto sociocognitivo é de trabalho profissional com ordenações e finalidades distintas. Por isso, a investigação do processo de leitura documentária, como ato cognitivo ou semiótico, precisa ir além e combinar as variáveis para descobrir como a leitura de um texto narrativo poderá

ser realizada a partir da finalidade de um sistema de recuperação de uma biblioteca universitária.

Assim, para atender diferentes lacunas que acreditamos existir, convidamos autores da área de Ciência da Informação que investigam diferentes finalidades de leitura documentária para diferentes estruturas textuais em diferentes contextos para compor esta coletânea buscando aprofundar o tema “leitura documentária”, ampliar a literatura produzida e buscar ampliar a reflexão sobre esta prática.

Com os resultados das investigações realizadas em torno da leitura documentária obtivemos os textos dos 11 capítulos desse livro que, além de revisar a literatura científica sobre o tema em Ciência da Informação e de áreas interdisciplinares, demonstram avanços e inovações com a finalidade de contribuir com atualização teórica e a elaboração de metodologias aos processos de Organização e Representação do Conhecimento

No primeiro capítulo temos o enfoque da leitura documentária e o processo de compreensão do indexador, um relato do memorial de investigação científica da Professora Mariângela Spotti Lopes Fujita. Este texto é não apenas enriquecedor em seu aporte teórico, mas também como um relato de pesquisa que nos auxilia em seu percurso metodológico. Ao final, traz resultados importantes obtidos de 6 projetos de pesquisa no período de mais de uma década de investigações: as “Diretrizes teórico-metodológicas sobre leitura documentária para indexação” e o “Modelo de leitura para indexação de textos científicos: manual explicativo” que, certamente, serão de muita valia para professores, pesquisadores, estudantes e profissionais.

O capítulo seguinte “Os processos cognitivos na leitura documental” traz reflexões teóricas para responder a pergunta “o que faz o indexador quando lê?” durante a análise de assunto para indexação. Com uma redação que envolve o leitor, este capítulo, tem abordagem cognitva para apresentar as operações que ocorrem no cérebro de um leitor e, assim, fazer uma análise dos processos cognitivos (microprocessos, macroprocessos, processos de integração, de elaboração e metacognitivos) durante a leitura documental. Ao final, apresenta a compreensão de leitura em ambiente de aprendizagem.

Com enfoque mais voltado às estratégias do profissional documentalista, o terceiro capítulo, descreve as estratégias de leitura documentária para indexação ao examinar, com uso da técnica introspectiva do Protocolo Verbal, os indexadores e o contexto de um sistema de informação especializado, o Centro de Informações Nucleares da Comissão Nacional de Energia Nuclear (CIN/CNEN). A principal conclusão do estudo é a proficiência do indexador em leitura documentária da documentação especializada em energia nuclear e áreas correlatas. O estudo contribuiu com a observação e demonstração das estratégias cognitivas e metacognitivas desse indexador proficiente.

Na perspectiva do documento de arquivo, o quarto capítulo investiga as convergências e divergências entre a análise documental no âmbito arquivístico e a representação descritiva e temática da informação no âmbito da Organização e Representação do Conhecimento. Considerando que tanto a análise documental quanto a representação temática e descritiva necessitam da leitura documentária, o estudo traz à tona pontos em comuns e incomuns nos diferentes âmbitos que influenciam o trabalho do documentalista.

O estudo apresentado no quinto capítulo investiga as possíveis orientações sobre leitura documentária no modelo *Functional Requirements for Subject Authority Data* (FRSAD) da *International Federation of Library Associations and Institutions* (IFLA). O objetivo é descobrir a relação da leitura documentária para a representação de assunto na construção de catálogos por meio do modelo, tendo em vista que designa, em nível conceitual, os aspectos referentes às entidades que representam atributos dos assuntos dos documentos de interesse para os usuários.

O tema indexação de literatura infantil do gênero fábula do sexto capítulo tem como variável a estrutura textual narrativa ficcional do gênero fábula infantil investigada pela leitura documentária. Com fundamento no Modelo de Leitura Documentária para indexação de textos científicos o estudo teve como resultado a proposta do Modelo de Leitura Documentária para identificação de conceitos ficcionais da literatura infantil do gênero discursivo fábula infantil com uso de estratégias de inferência ao texto por questionamento para a identificação de conceitos mediante conhecimento da superestrutura e macroestrutura textual.

O trabalho apresentado no capítulo 7 investiga recursos auxiliares semiautomáticos de indexação de textos com o objetivo principal de verificação de aplicabilidade dos padrões de Hearst no contexto da leitura documentária. Para observar, de modo semiautomático, a identificação de termos e suas relações hierárquicas, foi realizado experimento de aplicação do algoritmo de Hearst com uso do software Word Smith Tools, versão free 4.0. para a modalidade de análise semiautomática comparada à análise manual de artigos do fascículo mais recente da Revista Latino-Americana de Enfermagem. Os resultados obtidos demonstram, por um lado, a complexidade da leitura documentária ao se realizar a extração semiautomática de padrões léxico-sintáticos e, por outro lado, sua utilidade como recurso semiautomático aplicado ao crescente acréscimo de textos em formato eletrônico e os hipertextos possibilitarem o tratamento semiautomático de conceitos em função de seus formatos, estruturas e suportes.

Quando a leitura documentária se refere a documentos textuais com gráficos, tabelas e infográficos precisamos compreender a vizualização da informação. Essa informação visual é complementar ao texto do documento, porém, merecedora de atenção por revelar aspectos não ditos ou que complementam a idéia do texto, principalmente na comunicação jornalística e em publicações digitais na Internet que utilizam técnicas e ferramentas para representar a informação e ajudar na compreensão da leitura que cada vez mais se interessa em aportar gráficos e infografias que sintetizam idéias e conceitos. O capítulo 8 desenvolve estudo sobre a comunicação gráfica e a tipologia de visualização da informação presentes na comunicação científica que contribuem para a compreensão das investigações científicas dedicadas ao aprimoramento da representação do conteúdo textual durante a leitura documentária.

Os três capítulos (9, 10 e 11), que fecham a proposta deste livro, trazem a abordagem da Semiótica nos estudos teóricos sobre leitura documentária. O capítulo 9 que se refere à semiótica documental considera a indexação de assunto como processo semiótico e a leitura como etapa anterior determinante para os resultados de representação para a recuperação da informação. A questão a ser respondida por esse estudo do capítulo 9 é: qual disciplina, campo ou área dá suporte e abriga

contribuições para a compreensão do processo de leitura documental. Com o objetivo de explicar a importância epistemológica da Semiótica Documental para o acréscimo de elementos conceituais à leitura documental, o estudo apresenta a Linguística Documental e depois a idéia da Semiótica Documental, considerada pelo autor, como capaz de abrigar as variáveis da leitura documentária.

O capítulo seguinte propõe categorias de análise identificadas em contribuições de estudos semióticos aplicadas à análise de protocolos verbais coletados durante a leitura documental de catalogadores de bibliotecas universitárias para indexação de livros. As categorias de análise selecionadas com base em estudos teóricos são a semiose, níveis de interpretante, experiência colateral, estado de dúvida/crença e hábito/mudança de hábito. Essas categorias foram utilizadas para analisar trechos das transcrições de protocolos verbais que resultaram na discussão da leitura documental na perspectiva da teoria proposta por Peirce. As discussões teóricas que permeiam a análise da prática de leitura documentária por profissionais ressaltam a conclusão de que o processo de indexação constitui-se de interpretação e representação, conceitos relevantes na teoria semiótica proposta por Peirce.

Para fechar a tríade de trabalhos com abordagem semiótica o capítulo 11 trata das contribuições teóricas da semiótica para a leitura documentária. Por meio da visão de Peirce, a teoria semiótica explica três tipos de raciocínio, Abdução, Dedução e Indução, que podem esclarecer o processo de leitura documentária. O estudo apresentado no capítulo coloca em discussão comparada os três raciocínios e demonstra, com trechos de Protocolos Verbais de catalogadores, que fazem parte de um processo inferencial de indexação realizado em etapas seqüenciais de raciocínio do indexador observados na leitura documentária.

Buscamos, dessa forma, estender o conhecimento referencial e de inovação sobre leitura documentária para a motivação de novas pesquisas com resultados para os leitores do sistema de informação e leitores documentários dos mais diferentes contextos.

Referências

NEVES, D. A. de B. *Aspectos metacognitivos da leitura do indexador*. 131 f. Tese (Doutorado em Ciência da Informação) – Universidade Federal de Minas Gerais, Belo Horizonte, 2004. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/EARM-73FMVG/doutorado___dulce_am_lia_de_brito_neves.pdf?sequence=1>. Acesso em 06 jun. 2017.

A leitura documentária e o processo de compreensão do indexador: memorial de investigação científica

Mariângela Spotti Lopes Fujita

A situação ideal para o processo de compreensão da leitura, segundo a teoria interacionista, é a indissociabilidade entre as três variáveis, texto, leitor e contexto, o que nos leva a considerar que as dificuldades da análise de assunto para a indexação devem ser analisadas a partir de cada variável e não somente do leitor/indexador. Dessa forma, o autor como emissor e o texto como mensagem escrita desempenham papel importante na transmissão da informação, pois é a crença na racionalidade do autor, na sua intenção de ser informativo dizendo algo coerente, que leva o leitor a interagir com o texto, realizando esforços para construir um significado viável.

Em função da leitura como processo comunicativo é preciso destacar o **princípio cooperativo** de Grice (1982), considerado como base de toda comunicação humana. Segundo esse princípio, o autor no momento da escrita deve ter em mente o princípio cooperativo para que o leitor possa compreender suas idéias, que estão representadas no texto, a fim de garantir que a leitura seja um ato comunicativo coerente. Segundo Grice, o princípio cooperativo depende de uma norma comportamental dividida no que denominou de máximas conversacionais, a saber: a da **quantidade** (seja suficientemente informativo), a da **qualida-**

de (afirme apenas o que acredita ser verdadeiro), a da **relação** (diga algo relevante) e a do **modo** (seja claro).

A variável texto representa as idéias do autor e pode influenciar na compreensão do leitor quando utiliza recursos apelativos que mexem com seu emocional, omite informações relevantes sobre o assunto, quando o texto está impresso em letras pequenas demais que dificultam a leitura, ou a escrita apresenta problemas como orações muito complexas ou curtas demais, ou ainda, incoerentes. Kato (1985b), citada por Cintra (1987), considera a legibilidade influente no processo interativo leitor/texto e a qualidade do texto como fator básico de legibilidade, acompanhado do conhecimento prévio do leitor e o tipo de estratégia que o texto exige. Na qualidade do texto aponta vários fatores que a promovem: manutenção do tema, correção gramatical, adequação lexical e estrutura do texto.

Vale ressaltar que o conhecimento lingüístico e textual deverá facilitar sobremaneira a escrita, para o autor, e a leitura, para o leitor. Portanto, na comunicação escrita, tanto o autor como o leitor devem conhecer os benefícios de coerência e coesão para que o texto seja compreensível (KATO, 1986 e KLEIMAN, 1989).

A coerência e a coesão são definidas por Beaugrande e Dressler (1981), citados por Koch (1989, p. 18), da seguinte maneira:

> [...] a coesão concerne ao modo como os componentes da superfície textual - isto é, as palavras e frases que compõem um texto – encontram-se conectadas entre si numa seqüência linear, por meio de dependências de ordem gramatical. [...] a coerência diz respeito ao modo como os componentes do universo textual, ou seja, os conceitos e relações subjacentes ao texto de superfície são mutuamente acessíveis e relevantes entre si, entrando numa configuração veiculadora de sentidos.

Koch (1989, p.19) completa as definições, esclarecendo que a coerência é a responsável pela continuidade de sentido no texto porque é resultado de "uma complexa rede de fatores de ordem lingüística, cognitiva e interacional."; por outro lado, a coesão garante a legibilidade do texto pela explicitação de tipos de relações estabelecidas entre os elementos lingüísticos e, em muitos tipos de textos, é um mecanismo de manifestação superficial da coerência.

A leitura é, evidentemente, a base para o aprendizado da linguagem escrita, principalmente quando o leitor reconhece o tipo de texto por meio de sua estrutura. Cada texto possui suas próprias convenções de apresentação, tipografia e estilo que os torna distinto de outros. Assim, livros-textos não possuem os mesmos esquemas que os artigos de jornais, poemas, cartas etc. Essas distintas características de estruturas para a organização do conteúdo textual fornecem importantes subsídios ao conhecimento de leitores e escritores, facilitando a previsão. Quanto mais o leitor se familiariza com diferentes tipos de texto, mais experiente e hábil se torna para ler variadas espécies de textos.

Com a afirmação de que é necessário o conhecimento lingüístico e textual, tanto para o autor quanto para o leitor, a fim de se garantir a codificação e decodificação da mensagem transmitida no processo de comunicação pela leitura, estamos adicionando à compreensão o recurso da previsão. Conforme Smith (1989, p.35), "previsão e compreensão podem estar interligadas". De forma simples e genérica, previsão é a elaboração de perguntas sobre os assuntos que o leitor pretende encontrar no texto, possível pelo conhecimento da estrutura textual. Compreensão constitui as respostas às perguntas feitas na previsão.

Se a previsão leva à compreensão, isso significa que a leitura é uma atividade que envolve o pensar simultâneo, pois o leitor utiliza seu conhecimento prévio sobre o assunto para melhor compreender o texto. O pensamento pode ocorrer após a leitura, neste caso o leitor realiza uma reflexão subseqüente sobre o texto.

É importante considerar, além do recurso de compreensão e previsão, que a leitura pode apresentar quatro características primordiais - ser objetiva, seletiva, antecipatória e baseada na compreensão (SMITH, 1989):

- é objetiva porque as pessoas lêem com a finalidade de encontrar algo; esta característica de leitura é central porque o leitor antecipa a sua compreensão por meio das suas intenções.
- é seletiva porque o assunto que lhe interessa é colocado em evidência, ou seja, despreza-se o que é considerado redundante.

- é antecipatória porque o leitor tem um objetivo estabelecido e suas expectativas são definidas a partir dele, isto é, sabe o que pode encontrar no texto.
- é baseada na compreensão, ou seja, o leitor deve eliminar as dúvidas quanto à ambigüidade referente ao conteúdo ou palavras.

O conhecimento prévio para a compreensão em leitura depende do conhecimento existente na memória a longo prazo, um repositório de conhecimentos com tempo e capacidade de armazenagem permanente e ilimitado e que possui uma estrutura de conhecimento baseada em uma rede semântica de informações que liga seus "nós" mediante associações significativas entre conceitos, fatos, ações, etc., ali representados. Para se realizar o processo de compreensão é preciso, também, que a memória a longo prazo tenha os chamados "esquemas" ou representações generalizadas de ambientes, situações familiares e informações para que se faça associação com tudo aquilo que se está vendo, ouvindo e lendo.

Esquemas são "estruturas abstratas, construídas pelo próprio indivíduo, para representar a sua teoria do mundo. Na interação com o meio, o indivíduo vai percebendo que determinadas experiências apresentam características comuns com outras." (LEFFA, 1996, p.35). Na mente humana, tudo vai sendo transformado a partir do que o indivíduo já possui e, assim, os esquemas vão sendo incorporados na estrutura cognitiva para entrelaçar-se com outros esquemas. Para isso, os esquemas possuem variáveis que se associam aos diferentes aspectos de uma situação ou problema, como, por exemplo, o ato de vestir-se pode associar variáveis de temperatura, cor, finalidade da atividade para a escolha de uma roupa. Essas variáveis são subesquemas que constituem os esquemas. A cada vez que acionamos um esquema, podemos acionar subesquemas ou esquemas superordenados que comportam esquemas com uma mesma característica.

Os "esquemas" são objeto de estudo da área de cognição que os vincula ao conhecimento prévio armazenado na memória a longo prazo porque "são considerados como representações de padrões ou regularidades mais gerais que ocorrem em nossa experiência." (SMITH, 1989, p.30). Isso significa que o esquema existe por nosso conhecimento prévio e nos

dá condições de prever ou de antecipar atividades, acontecimentos, ações e informações ao considerar o certamente provável e não o improvável.

Assim, um texto sob a forma de um artigo científico será mais rapidamente compreendido se o leitor já souber como é a organização característica deste tipo de texto e estiver familiarizado com a estrutura tradicional. Essa capacidade de predição aumentará à medida que o leitor aumentar a sua compreensão, sobrepondo informações umas às outras e aumentando sua estrutura de conhecimento.

Quando lemos algum texto, sempre estamos prevendo encontrar algum aspecto ou padrão compatível aos nossos esquemas que nos ajude a compreendê-lo.

> Sem o acionamento de um esquema, a compreensão não é possível. Ao iniciar a leitura de um texto, a primeira coisa que o leitor normalmente faz é vasculhar a memória em busca de um esquema em que ele possa fixar as informações do texto (LEFFA, 1996, p.38).

A teoria de esquemas foi elaborada a partir de várias pesquisas que investigaram o processo de leitura e desenvolveram modelos classificados a partir de três critérios (PINTO e GÁLVEZ, [1996], p.42):

> 1) Segundo sua estrutura:
>
> - modelos seriais (Foster, 1979), (Garret, 1975), (Mitchell, 1982);
>
> - modelos em paralelo (Marslen-Wilson, 1985), (Dell & Reich, 1981), Stemberger, 1985);
>
> 2) Segundo a direção do fluxo de informação:
>
> - Modelos ascendentes - bottom-up: (Gough, 1972), (Laberge & Samuels, 1974)
>
> - Modelos descendentes - top-down: (Goodman, 1971)
>
> - Modelos interativos: (Rumelhart, 1977), (Vega et al., 1990)
>
> 3) Segundo seu funcionamento:
>
> - modelos modulares: (Fodor, 1983)
>
> - modelos globais: (Anderson, 1983)

A leitura, segundo o modelo interativo de Rumelhart, é um ato duplo: de recepção ou percepção visual e de compreensão ou atividade mental (PINTO MOLINA,1993, p.161). Isso significa que ao ler, o leitor realiza dois movimentos inversos e, ao mesmo tempo, complementares: *"bottom-up"*, ascendente ou indutivo e "*top-down*", descendente ou dedutivo.

Kato (1985a, p.82) indica que a noção de "esquemas" (schemata), proposta por Rumelhart e Ortony (1977) e retomada por Rumelhart (1977), é considerada uma proposta atrativa por pretender ser "...uma teoria do conhecimento que engloba uma *teoria prototípica* do significado e ser, ao mesmo tempo, uma *teoria de procedimento*...", o que significa que os movimentos *bottom-up* e *top-down* são processos necessários para a compreensão.

Durante o processo de leitura, o leitor pode ativar seus "esquemas" mediante dois movimentos complementares:

> **movimento botton-up**, em que o leitor vai lendo na dependência do contexto escrito, ou seja, vai extraindo, linearmente, dos símbolos impressos o significado, caminhando das partes para o todo, e **movimento top-down**, no qual há maior dependência de conhecimento prévio do leitor, pois ele vai fazendo generalizações e predições a partir de "**esquemas**" que tem armazenados em sua memória, formulando hipóteses que ajudarão na compreensão do texto. (CINTRA, 1987, p.31)

Kato (1985a) considera leitor experiente aquele que utiliza os dois tipos de estratégias, as ascendentes (dependentes do texto, da análise cuidadosa do *Input* visual) e as descendentes (baseadas no conhecimento prévio do leitor e na sua capacidade de inferência, de predição), relacionando ora um tipo, ora outro, de maneira consciente, no momento em que cada uma delas se fizer necessária. Há momentos na leitura em que um trecho difícil para o leitor exige que ele leia linear e cuidadosamente, e há outros em que apenas inferências contextuais permitem a compreensão sem problemas.

Acreditamos que o leitor estratégico e proficiente seja aquele que, além de utilizar apropriadamente estratégias ascendentes e as descendentes, mantém em mente o objetivo da leitura.

Kato (1985a, p. 87) afirma que os esquemas, além de regras e taxonomias, contêm um componente de procedimentos (estratégias) que dita como utilizar suas partes (esquemas para subesquemas e superesquemas).

De acordo com Nardi (1993, p. 25), a ativação de subesquemas e superesquemas determina o acionamento dos movimentos *top-down* e *bottom-up*, ou seja, "Se na interpretação de um evento, a ativação de um esquema, ativa seus subesquemas, dizemos que houve processamento *top--down* (do todo para as partes); se por outro lado, ativar superesquemas, o processamento terá sido *bottom-up* (das partes para o todo)."

Kato (1985a, p.83) esclarece que "...a adivinhação é parte da estratégia *top-down*, por ser esta mais preditiva, mas a *bottom-up* seria responsável pela confirmação, pelo refinamento..."

A característica principal deste modelo do processo de leitura de Rumelhart é a interação dos componentes de cada nível, a ativação de todos os subprocessos e a influência da informação contextual em todo o processamento.

Pinto e Gálvez ([1996], p. 45) consideram o modelo interativo de Rumelhart (1977) é o que melhor representa a atuação do sujeito documentalista[1] porque:

> [...] quando está lendo um texto cujo assunto lhe seja desconhecido, procederá dos níveis mais inferiores ao superiores, fará uma leitura lenta, detalhada, ascendente ou bottom-up. Ao contrário, quando o texto for familiar realizará uma leitura "entre linhas" - descendente ou top-down -, antecipando informações e dirigindo-se a uma representação do conteúdo global do texto.

De outra forma, os movimentos ascendentes e descendentes do processo de leitura, a partir do modelo de Rumelhart, são diretamente relacionados a duas definições opostas que esclarecem melhor a função de cada movimento na leitura (LEFFA, 1996, p.11):

- Ler é extrair significado do texto;
- Ler é atribuir significado ao texto.

[1] Documentalista para os teóricos espanhóis significa o mesmo que indexador em ciência da informação no Brasil

Na primeira afirmação, o movimento é *bottom-up* e a leitura é um processo ascendente. A compreensão sobe do texto ao leitor na medida exata em que o leitor vai avançando no texto. Na segunda afirmação, o movimento é *top-down* e a leitura é um processo descendente, pois desce do leitor ao texto. Enquanto vai lendo, o leitor prevê hipóteses sobre o que irá tratar o documento e, no decorrer da leitura, suas hipóteses podem ou não ser confirmadas. As hipóteses serão elaboradas de acordo com o conhecimento prévio do leitor (LEFFA, 1996, p.13).

A diferença entre as duas concepções de leitura está justamente na interação, pois no processo ascendente, não existirá uma interação entre o leitor e o texto, o leitor lê o texto linearmente, palavra por palavra, extraindo o seu significado de acordo com a seqüência lingüística apresentada pela estrutura do texto, enquanto que, no processo descendente, a obtenção do significado do texto se dá por meio da contribuição do leitor com suas hipóteses e previsões.

Segundo Kato (1985a, p.74), enquanto o movimento ascendente (*bottom-up*) é usado para designar os processos lineares, sintéticos e indutivos, o movimento descendente (*top-down*) é usado para o processo não linear, analítico e dedutivo.

É preciso, entretanto, reconsiderar que durante a leitura os movimentos ascendentes e descendentes poderão se alternar, daí a prevalência do modelo interativo de leitura de Rumelhart (1977), pois os movimentos são complementares.

Em estudo de John Farrow (1991), observa-se que os movimentos *bottom-up* e *top-down* são considerados, respectivamente, perceptual e conceitual e referem-se à leitura de indexadores e resumidores.

Farrow (1991, p. 153), baseado em estudos de observação de Masson (1982) com leitores rápidos (*skimmers*) que realizaram processamento perceptual (*bottom-up*) e conceitual (*top-down*), considera que o método de processamento de texto de indexadores e resumidores é semelhante ao processo perceptivo, com o qual os *skimmers* procuram dicas, palavras em itálico, utilizando seu conhecimento de estrutura textual para fixar o olhar em áreas relevantes do texto. Porém os estudos de Masson, segundo Farrow (1991), também indicam

que os leitores rápidos realizaram um processo seletivo de informação essencial baseado mais no processo conceitual do que perceptivo porque tinham um objetivo específico na leitura. No caso de indexadores e resumidores, Farrow (1991, p.155) aponta que, apesar da indexação possuir um objetivo específico, não existe nenhuma evidência de que os sujeitos realizaram um processo seletivo da informação essencial baseado amplamente em processamento conceitual a partir de um objetivo específico.

Tais modelos de compreensão, baseados no conhecimento prévio, são considerados interativos no que diz respeito à interação entre os processos *bottom-up* e *top-down*, bem como do leitor com o texto, mas não à interação do leitor com o autor (KLEIMAN, 1989, p.31).

Deste ponto de vista, a leitura é um processo interativo e comunicativo que se dá entre o leitor (determinado pelo seu contexto) e o autor por meio do texto. Segundo Kleiman (1989, p.33), dentro de uma perspectiva interacionista, as relações entre o leitor e o texto estabelecem, também, uma relação daquele com o autor.

Reportando-se ao princípio cooperativo de Grice (1982), Cintra (1987, p.29) considera que na leitura para fins documentários o princípio de cooperação autor/leitor é rompido porque o autor não previu o documentalista como leitor. Considerando que o indexador é, em princípio, um leitor, acredito que não foram previstos seus objetivos de análise, diferentes de um leitor que busca conhecimento ou informação, pois o indexador, de acordo com seus objetivos e contexto, é também um leitor que compreende o texto para interpretação e produção de uma representação condensada de seu significado.

Ao avançarmos na concepção de esquemas, podemos deduzir que o indexador, como leitor, terá o conhecimento lingüístico prévio, implícito e importante para a compreensão da organização textual. Portanto, na "...leitura para fins documentários o leitor não lê palavra por palavra, letra por letra, e muitas vezes não lê freqüentemente todas as palavras do texto, quando as seqüências são previsíveis." (CINTRA, 1987, p. 30). Ao contrário, o leitor indexador saltará trechos em busca da informação relevante.

Pelo processo *top-down*, o leitor realiza inferências ao conteúdo textual para deduzir e predizer, baseando-se em seus esquemas que incluem um quadro de referências composto de entidades lingüísticas e conceituais.

A construção do sentido do texto, segundo Kleiman (1989, p. 13), depende da "...interação de diversos níveis de conhecimento, como o conhecimento lingüístico, o textual e o conhecimento de mundo." Segundo a autora, o conhecimento textual desempenha um papel importante na compreensão de textos.

Cintra (1987, p.31), baseada em pesquisas que demonstram a facilidade de leitura para leitores com conhecimento de estruturas textuais, considera que "o leitor que domina as superestruturas textuais capta com mais facilidade as idéias centrais do texto, pois tem como parâmetro a identificação dos constituintes básicos."

Para o indexador, o domínio da tipologia documentária e da estrutura textual são dois tipos de conhecimentos prévios que poderão aumentar sua compreensão durante o processo descendente de leitura. Segundo Ginez de Lara (1993, p.55), "...envolve, portanto, tanto o reconhecimento da tipologia textual, como a identificação dos elementos referenciais para uma interpretação apropriada."

De acordo com Giasson (1993, p.25), "o leitor aborda a atividade de leitura com as estruturas cognitivas e afetivas que lhe são próprias. Além disso recorre a diferentes processos que lhe permitem compreender o texto."

Os processos de leitura são utilizados pelo leitor como habilidades para abordagem do texto durante o desenvolvimentos das atividades de cognição. Existem diferentes tipos de processos para cada ação de compreensão que se realizam em diferentes níveis e simultaneamente. Segundo Giasson (1993), existem processos para a compreensão da frase (microprocessos - nível da frase); para a busca de coerência entre as frases (processos de integração - entre as frases); para construir um modelo mental do texto (macroprocessos - nível do texto); para permitir ao leitor captar os elementos essenciais e levantar hipóteses (processos de elaboração - extensão do texto) e; para gerir a compreensão (processos metacognitivos - compreensão do texto).

A interação entre o texto e o leitor desenvolve-se pelo uso de estratégias, definidas. As estratégias de leitura, ou as ações que o leitor realiza no ato de ler, têm sido definidas por vários autores. Essas estratégias, segundo Faerch e Kasper (1980), citados por Nardi, (1993), são planos potencialmente conscientes do leitor para resolver algo que se apresenta como um problema na compreensão. Brown (1980, p.456) define estratégia como "qualquer controle deliberado e planejado de atividades que levam a compreensão". Para Oxford (1989, apud Nardi, 1993, p.19), estratégias "são ações direcionadas para um objetivo, potencialmente observáveis, potencialmente ensináveis e flexíveis". Para a autora, as estratégias não podem ser prontamente observáveis e sim as ações comportamentais do leitor (como, por exemplo, o virar de páginas, ou a procura de uma palavra no dicionário), mas as ações mentais como associações e deduções durante a leitura não podem ser vistas.

Contudo, o ponto de contradição entre os autores é o grau de consciência. Brown (1980) acredita que as ações são intencionalmente selecionadas, enquanto Faerch e Kasper (1980) referem-se a ações potencialmente conscientes. Por isso, autores seguidores de Brown, fazem distinção entre estratégias e habilidades automáticas. Palincsar e Brown (1984) acreditam que leitores proficientes freqüentemente usam "Skills", e consideram estratégias apenas o comportamento metacognitivo (consciente) frente a um problema. Com base em Brown (1980), Nardi (1993, p.20) distingue "Skill" de "Estratégia" com a explicação de que:

> Skill seria uma estratégia que teria sido adquirida em algum momento da aprendizagem e se tornado automática (raramente "recuperada" pelo leitor proficiente), e estratégia seria o uso consciente de uma "Skill" ou de uma nova "tática", em momentos de solução de problemas.

Kato (1985a) distingue dois tipos de estratégias que definem o comportamento do leitor: as estratégias cognitivas, que são aquelas automáticas e subconscientes, utilizadas durante a leitura fluida, sem obstáculos, e as estratégias metacognitivas, que são ações conscientes do leitor frente a um problema.

As estratégias cognitivas de Kato (1985a) são denominadas, por Cavalcanti (1989), estratégias automática, e as metacognitivas são denominadas estratégias controladas.

Como visto anteriormente, o uso de estratégias não é facilmente observável porque ações mentais, como associações e deduções durante a leitura, não podem ser vistas, ainda que, possam ser verbalizadas.

Para conferir natureza metacognitiva às ações mentais, Brown (1980, p.456) indica as seguintes atividades:

- explicitação dos objetivos da leitura;
- identificação de aspectos importantes da mensagem;
- alocamento de atenção a áreas importantes;
- monitoração do comportamento para ver se está ocorrendo compreensão;
- engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido;
- tomada de ações corretivas quando são detectadas falhas na compreensão;
- recobramento de atenção quando a mente se distrai ou faz digressões.

O critério geralmente usado para separar as atividades cognitivas das metacognitivas é o do envolvimento da consciência: as atividades cognitivas estariam abaixo do nível da consciência; as metacognitivas envolveriam uma introspecção consciente (BROWN, 1980, p.455)

A metacognição em leitura permite ao leitor uma compreensão de sua própria compreensão, ou melhor, o acompanhamento e avaliação de seu processo de compreensão durante a leitura de um texto e, além disso, a tomada de providências quando a compreensão falha (LEFFA, 1996, p.45).

Na classificação das atividades cognitivas e metacognitivas, Leffa (1996, p.49) propõe que sejam diferenciadas não apenas pelo critério de envolvimento da consciência, mas também pelo tipo de conhecimento

utilizado para executar a atividade, seja declarativo ou processual. Segundo Tardiff (1997, p.186), na interação com um texto, o leitor trata as informações baseando-se em seus conhecimentos declarativos, assim como em seus conhecimentos condicionais e de procedimento. Tardiff (1997) acrescenta o conhecimento condicional, junto ao conhecimento processual, ambos integrados às estratégias metacognitivas, porque correspondem ao "quando agir", o conhecimento da estratégia de leitura adequada a ser usada no momento certo e "por que agir", o conhecimento que justifica as razões pelo uso de determinada estratégia.

Conforme Leffa (1996, p.49), o conhecimento declarativo faz parte das atividades cognitivas envolvendo a consciência apenas da tarefa a ser executada, ou seja, o leitor usa o conhecimento para realizar, por exemplo, a indexação de um documento e a realiza porque sabe que é capaz, ou mesmo para ler um livro, pois tem conhecimento que sabe ler. Com o conhecimento processual, porém, a consciência vai além da tarefa a ser executada, o leitor tem a consciência de sua própria consciência, ou seja, tem controle do próprio conhecimento e do processo que deve seguir para atingir o resultado. Neste sentido, o conhecimento declarativo, e vale a pena entender, pelo exemplo de Leffa (1996, p.49-50), engloba o conhecimento declarativo pela atividade cognitiva e o conhecimento processual pela atividade metacognitiva,

> Um exemplo de consciência do processo pode ocorrer, por exemplo, na leitura de um romance em que o leitor está absorvido pelos acontecimentos narrados pelo autor. A leitura vai fluindo rápido, os processos ascendentes e descendentes trabalhando numa orquestração perfeita. O leitor não tem nenhuma consciência do processo da leitura, mas concentra toda sua atenção no efeito que obtém da leitura. Tem consciência do que o personagem principal está fazendo, mas não tem consciência de sua própria leitura; não se dá conta, por exemplo, se está lendo rápido ou devagar. De repente, porém, pode surgir um problema. O que o leitor pensava ser a fala de um personagem é a fala de um outro, que não consegue mais identificar. O leitor dá-se conta de que não está mais compreendendo o texto. A leitura que seguia fluida e rápida é bruscamente interrompida e o leitor decide voltar algumas linhas para poder retomar o fio do enredo. A leitura rápida e fácil, concentrada no conteúdo, é uma atividade cognitiva. A descoberta de que houve um problema e de que uma correção no rumo da leitura tinha que ser feita para recuperar o texto é uma atividade metacognitiva.

O conhecimento processual permite que a leitura seja consciente, que o leitor perceba a forma como o texto está sendo lido e o nível de compreensão que está sendo atingido por ele. Cavalcanti (1989) considera que as estratégias se tornam mais observáveis quando ocorre algum tipo de ruptura na compreensão, momento em que o leitor deverá desacelerar a leitura e tornar-se metacognitivo. Essa ruptura pode ser causada por um déficit em algum dos componentes lingüísticos da competência comunicativa. Portanto, é possível observar o processo de leitura pela verbalização do conhecimento processual do leitor por meio de Protocolo Verbal.

O uso de estratégias cognitivas e metacognitivas certamente deverá tender a um equilíbrio, pois, segundo Cintra (1987, p.32), "Ainda que toda leitura envolva esses dois tipos de estratégias, é provável que quanto menos atividades metacognitivas exigir, mais legível será o texto. Entretanto, é também provável que a leitura apenas automática conduz à incompreensão".

Os dados obtidos em pesquisas feitas sobre a metacognição da leitura sugerem quatro conclusões principais (LEFFA, 1996, p.64):

- A metacognição desenvolve-se com a idade;
- A metacognição correlaciona-se com a proficiência em leitura: leitores fluentes têm mais consciência de seus comportamentos de leitura e possuem maior flexibilidade para ajustar a leitura a objetivos específicos;
- O comportamento metacognitivo melhora com a instrução: o treinamento específico das habilidades metacognitivas tem feito o aluno responder de modo mais eficaz a mensagens ambíguas do falante;
- A eficácia de uma determinada estratégia depende do objetivo de uma determinada leitura. Em termos absolutos, as estratégias que consomem mais tempo, tais como reler ou sublinhar palavras-chave, são as que levam a uma compreensão mais profunda e crítica de um texto.

Kato (1985b) considera que as estratégias cognitivas são regidas pelos princípios de Canonicidade e Coerência. O princípio da Canonicidade significa que o leitor possui conhecimento da ordem natural sintática e semântica permitindo-lhe predizer, por exemplo, a categoria gramatical de uma palavra desconhecida e assim facilitando a inferência de seu significado.

Para a tarefa do indexador esses princípios são importantes, considerando-se que um texto sob a forma de artigo científico, por exemplo, possui uma estrutura de tópicos e parágrafos já conhecida e o indexador possua habilidade em indexação de uma determinada área de assunto, a legibilidade do referido texto aumenta, assim como o uso de estratégias cognitivas.

O leitor deve, então, buscar detectar a estrutura do texto, pois o reconhecimento das superestruturas textuais favorece a captação das idéias principais do texto e seus conhecimentos prévios, a inferência de significados e levantamento de hipóteses que o ajudarão a apreender a temática global.

A esse respeito, Cintra (1987), em seu estudo sobre estratégias de leitura em documentação, apóia-se em Kato (1985b) para explicar que a legibilidade da leitura depende de três fatores básicos: a qualidade do texto, o conhecimento prévio do leitor e o tipo de estratégias que o texto exige. Apesar de o primeiro fator, conforme explica Cintra (1987), não depender do indexador, os outros dois estão, de alguma forma, vinculados à habilidade do indexador.

Entendemos que as atividades indicadas pela Norma 12.676 (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992) são metacognitivas no caso do leitor indexador ao identificarmos nelas aspectos análogos às atividades de natureza metacognitiva de acordo com Brown (1980). A análise demonstrada no QUADRO 1 (FUJITA, NARDI, SANTOS, 1998) pautou-se nas considerações teóricas realizadas em torno da natureza metacognitiva da leitura documentária, procurando identificar todos os aspectos da natureza das estratégias metacognitivas listadas por Brown (1980) - algumas reformuladas sob o ponto de vista de Kato (1985a) - e associadas às estratégias da Norma 12.676.

Atividades durante a leitura (BROWN, 1980; KATO, 1985a)	**Leitura Documentária (Norma 12.676)**
Explicitação dos objetivos da leitura e/ou manutenção dos objetivos na mente	
Identificação de aspectos importantes da mensagem	• Identificação de conceitos (abordagem sistemática mediante questionamento)
Alocamento de atenção a áreas importantes – Exploração da estrutura textual	• Análise do documento com domínio da estrutura textual, considerando partes do texto
Monitoração do comportamento por meio de: – Engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido – Tomada de ações corretivas quando são detectadas falhas na compreensão – Recobramento de atenção quando a mente se distrai ou faz digressões	
	Seleção de conceitos

Quadro 1: Identificação da natureza metacognitiva na leitura documentária
Fonte: FUJITA, NARDI, SANTOS, 1998

Considera-se que, sob o ponto de vista cognitivo da literatura revisionada, o indexador/classificador/resumidor são leitores aptos à compreensão pela própria estrutura cognitiva inata e construída: possuem conhecimento prévio, constituído de conhecimento lingüístico, textual, e conhecimento de mundo; utiliza seu conhecimento prévio por meio de esquemas acionados pelos movimentos *bottom-up* e *top-down* e realiza processos de compreensão, principalmente os metacognitivos.

A análise, quanto aos movimentos do leitor durante a leitura, propiciará a identificação de estratégias relacionadas à compreensão da leitura que envolvem a análise das atividades de cognição. A partir dessa análise é possível observar se o leitor indexador/classificador/resumidor utilizou estratégias de natureza metacognitiva ou cognitiva e assim determinar seu grau de proficiência para a leitura documentária que resultarão na melhor identificação de conceitos válidos para a recuperação do documento. Além disso, é possível observar procedimentos de análise de assunto e a seqüência adotada, visto que a leitura documentária é mais direcionada aos objetivos de indexação, classificação ou elaboração de resumos e, portanto, diferente de outras.

O indexador/classificador/resumidor é entendido como leitor profissional que realiza uma leitura documentária com objetivos definidos. O profissional da informação, visto como leitor é considerado individualmente na **abordagem cognitiva** pelo processamento de informações que realiza para **tratamento temático da informação documentária**. Contudo, como leitor profissional deve ser visto dentro de seu contexto sociocultural que abrange atuação e formação profissional em **abordagem sociocognitiva.**

A linha de investigação sobre leitura documentária desenvolveu, durante o período de 1993 a 2010, seis projetos de pesquisa com produção científica e formação de recursos humanos a partir dos eixos de abordagem cognitiva e abordagem sociocognitiva. No eixo de abordagem cognitiva a investigação, desenvolvida no período de 1993 a 2004, tinha como objeto de pesquisa a cognição do profissional da informação para observar seus procedimentos e estratégias de leitura de textos. Por outro lado, a investigação desenvolvida no período de 2004 a 2010, passou a explorar o contexto do leitor profissional no eixo da abordagem sociocognitiva. Em síntese, as variáveis de leitura, leitor e texto, foram investigadas nos projetos de pesquisa com abordagem cogntiva e a variável contexto nos projetos de abordagem sociocognitiva, conforme se apresenta, resumidamente abaixo, um cronograma dos projetos de pesquisa.

Projetos com abordagem cognitiva – Eixo I
1993 - 1995 - A organização do pensamento através da estruturação lógica do conhecimento: uma proposta de aplicação do Sistema de Indexação PRECIS para análise e compreensão literal de leitura
1996 - 2000 - Leitura em análise documentária
2000 - 2004 - Leitura em análise documentária: uma contribuição à formação do pesquisador
Projetos com abordagem sociocognitiva – Eixo II
2004 - 2007 - A leitura documentária na formação inicial do indexador: a abordagem sóciocognitiva na investigação de estratégias de ensino
2006 - 2009 - Política de tratamento da informação documentária da rede de bibliotecas da UNESP
2007 – 2010 - O contexto da leitura documentária de indexadores de bibliotecas universitárias em perspectiva sócio-cognitiva para a investigação de estratégias de ensino.

O estudo teórico e metodológico da leitura teve como objetivos:

- fornecer subsídios teórico-práticos para a compreensão do significado e dos procedimentos de leitura com fins de análise documentária; e
- realizar observação dos procedimentos de leitura documentária utilizados pelo indexador para exame do documento e identificação de conceitos.

A investigação sobre leitura documentária teve uma evolução natural que se iniciou com o projeto de aplicação da fundamentação teórica de indexação do sistema de indexação PRECIS, "A organização do pensamento através da estruturação lógica do conhecimento: uma proposta de aplicação do Sistema de Indexação PRECIS para análise e compreensão literal de leitura" (1993-1995), porque o PRECIS preconizava uma filosofia de preservação do contexto documentário durante e após a indexação com a finalidade de recuperação do assunto do documento e para isso indicava uma análise conceitual por meio do que denominava "conceitos universais", a ação, o objeto da ação, o agente que praticou a ação, tempo em que foi realizada a ação e da ação.

A temática desse projeto também se apresentava como continuidade, tendo em vista que o sistema PRECIS continuava sendo o fio condutor. O interesse partiu dos subsídios apresentados pela metodologia de análise conceitual do PRECIS para facilitar a análise de assunto de conteúdos de documentos por meio da compreensão de leitura, o que resultaria em uma organização lógica do conhecimento pelo leitor. Os resultados obtidos confirmaram que a análise conceitual do PRECIS é uma metodologia que facilita a compreensão de leitura e resulta em uma síntese objetiva do conteúdo do documento analisado. As conclusões propiciaram o desenvolvimento dos Projetos Integrados "Leitura em análise documentária" (1996-2000) (FUJITA, 1999b) e "Leitura em análise documentária: uma contribuição à formação do indexador" (2000-2004) e sua divulgação em artigo publicado no periódico Informare (FUJITA, 1999a) e em resumos de trabalhos apresentados em eventos nacionais e internacionais.

A análise conceitual do PRECIS funcionou como esquema de organização lógica do conhecimento para o leitor e, mais do que isso, é um dos elementos essenciais para a etapa de identificação e seleção de conceitos durante a leitura do texto com a finalidade de análise de assunto para indexação, classificação ou elaboração de resumos. A análise conceitual do PRECIS é compatível com o questionamento da Norma ABNT para a identificação dos conceitos. Dessa forma, a análise conceitual do PRECIS é um dos elementos fundamentais da metodologia de leitura documentária do Modelo de leitura documentária (Apêndice 1), criado durante o desenvolvimento do Projeto de Pesquisa "Leitura em análise documentária" e parte fundamental do "Programa de Orientação à formação e capacitação do indexador em leitura documentária" (Apêndice 2).

Com a evidência, obtida no relatório da pesquisa anterior, de que o leitor para atingir seu objetivo de leitura precisa compreender o conteúdo do texto e que tal processo de aquisição apresenta dificuldades sem uma metodologia de análise, passamos a inferir sobre as dificuldades de um indexador frente à análise de um texto com o objetivo de identificar conceitos representativos para futura recuperação pelos usuários. A principal premissa, portanto, era de que a leitura constitui atividade fundamental da Análise Documentária e, dessa perspectiva, estudos aprofundados sobre leitura poderão causar maior influência sobre o desempenho de indexadores e melhorar o uso de metodologias de indexação.

O Projeto "Leitura em Análise documentária" foi desenvolvido, no período de 1996 a 2000 em duas etapas: a primeira, no período de 1996 a 1998 e a segunda de 1998 até o ano 2000. Na primeira etapa foram realizados estudos teóricos sobre leitura e estratégias de leitura em Análise documentária e estudo de caso com bibliotecas universitárias brasileiras na área de Ciências da saúde oral.

O estudo formal da leitura foi importante no desenvolvimento da primeira etapa, porque, no contexto da análise, é a fase inicial que desencadeia o desempenho de todas as outras operações. A leitura em análise documentária é mais direcionada aos objetivos de indexação, diferente, portanto, de outras leituras. Assim sendo, o desenvolvimento do estudo sobre leitura em análise documentária direcionou seu foco de observação para o indexador, como agente que pratica a leitura, com a finalidade de

observar os procedimentos de leitura para exame do documento e identificação de conceitos. Com esses resultados foi possível a proposição do Modelo de Leitura Documentária (Apêndice 1) para identificação e seleção de conceitos durante a análise de assunto.

O estudo de caso analisado propiciou a caracterização do documentalista/indexador, do ambiente, das condições e variáveis em que atua, fez o levantamento das atividades de análise documentária para estudo dos procedimentos executados para a leitura documentária e, principalmente, realizou entrevista introspectiva e retrospectiva com Protocolo Verbal para observação da leitura dos indexadores, gravando o "Pensar Alto" para coleta e análise dos dados.

A observação do processo com Protocolo Verbal foi importante porque respaldou o depoimento do entrevistado e, para a pesquisa, tornou-se uma descrição ideal das estratégias de leitura.

Desenvolveram-se, assim, dois estudos caso para observação: o primeiro com bibliotecas universitárias na área de Odontologia, em São Paulo e interior, cooperantes da Sub-Rede Nacional de Informação em Ciências da Saúde Oral da BIREME, em que os sujeitos são bibliotecários e a tipologia documentária, o texto científico na área de Odontologia; e o segundo, no Centro de Informações Nucleares (CIN) da Comissão de Energia Nuclear, na cidade do Rio de Janeiro, em que os sujeitos são especialistas em energia nuclear sem formação em Biblioteconomia e a tipologia documentária do artigo de periódico científico na área de energia nuclear. Essas investigações foram desenvolvidas com a colaboração das instituições que tornaram disponíveis dados institucionais, para estudo do contexto físico, social e psicológico de indexação, bem como seus indexadores que se submeteram à metodologia do Protocolo Verbal.

A técnica introspectiva do Protocolo Verbal passou a ser utilizada pelas pesquisas em leitura documentária e trouxe resultados importantes da observação do pensamento de leitores proficientes e aprendizes durante a tarefa de indexação, elaboração de resumos e classificação que aprimoraram a metodologia de leitura documentária e a aprendizagem de leitores profissionais.

O desenvolvimento da segunda etapa do Projeto "Leitura em análise documentária" (1998/2000) investigou a fundamentação teórica sobre estrutura textual e identificação de conceitos e realizou estudo de caso com o Centro de Informações Nucleares - como outro serviço de análise - para observar e comparar resultados de estratégias de leitura quanto a diferentes objetivos: elaboração de resumos e indexação, em diferentes áreas de assunto (odontologia e energia nuclear).

O indexador, uma das variáveis influentes na leitura documentária, além do texto e do contexto, é também o ponto de partida para se obter um depoimento quanto aos fatores mais influentes na análise de assunto e sua concepção de análise. Com a análise dos protocolos verbais foi possível a obtenção de subsídios relativos à influência de toda a formação do indexador na leitura documentária e proporcionar contribuições para o Programa de Orientação à formação do indexador. Essa segunda análise dos protocolos verbais teve como parâmetro a classificação de Albrechtsen (1993) quanto às concepções de análise de assunto, caracterizadas conforme seus aspectos mais representativos devidamente esclarecidos em tabela realizada para essa finalidade.

A investigação procurou aprofundar a investigação em torno da análise de assunto para indexação e suas concepções de análise em busca de um referencial teórico. Obteve-se, no estudo de concepções de análise, a necessária confirmação de que a identificação e a seleção de conceitos são etapas imprescindíveis da análise de assunto e que são realizadas durante a leitura. Com a importância do processo de identificação de conceitos para a leitura do indexador realizou-se revisão de literatura que considerou os principais estudiosos da área bem como as normativas e manuais de serviços de análise expondo a discussão a respeito da preservação da tematicidade inerente do documento durante a identificação de conceitos. Pela subjetividade que envolve a análise de assunto, tomou-se conhecimento das teorias sobre a determinação do assunto, operação delicada da análise de assunto e abordada em estudos teóricos, ora sobre o conceito de assunto ("subject") ora sobre tematicidade ("aboutness"). Indiscutivelmente, a determinação de assunto é a etapa crucial em que o indexador como leitor é a principal variável, ainda que sujeito às suas concepções de análise de assunto e à política de indexação do sistema de informação.

A pesquisa da segunda etapa foi complementada, ainda, com o estudo de estratégias de leitura em diferentes estruturas textuais: texto jornalístico (utilizando o arquivo do jornal "O Estado de S. Paulo" como outro estudo de caso) e, também, o estudo de avaliação de eficácia de indexação e de resumos mediante uso de linguagens documentárias, tendo em vista a observação de resultados tanto a respeito da 'associação com linguagem' durante a leitura do indexador de literatura odontológica, quanto a respeito da compatibilidade dos termos com descritores da linguagem.

Na seqüencia foram realizados os projetos "Leitura em análise documentária de artigos de jornal" e "Avaliação de linguagem documentária em catálogos cooperativos *online*: um estudo de caso para levantamento de indicadores de avaliação do banco de dados bibliográficos ATHENA", com desenvolvimento vinculado ao Projeto "Leitura em análise documentária".

O estudo teórico e metodológico da leitura desenvolveu-se com a revisão de literatura em leitura e leitura documentária para reunir subsídios teóricos sobre as condições da leitura e os procedimentos de leitura documentária em indexação, suficientes para entender o indexador como leitor.

No estudo da literatura teórica, muitas investigações foram realizadas em torno da leitura e da leitura documentária, analisando seus aspectos cognitivos e lingüísticos com especial enfoque para o indexador como leitor e sua interação com o texto e seu contexto, tendo em vista a abordagem teórica e metodológica dentro da perspectiva interativa de Cavalcanti (1989) e Giasson (1993).

O aspecto cognitivo da leitura teve embasamento principalmente em Brown (1980), Kato (1985a, 1986), Cavalcanti (1989), Kleiman (1989, 2000), Giasson (1993), e Nardi (1993) além de abordar a teoria de esquemas de Rumelhart (1977). Em leitura documentária foi realizada dupla abordagem: cognitiva e lingüística. Na abordagem cognitiva a base teórica foi formada com Cintra (1987), Farrow (1991,1995,1996), Pinto Molina (1993) e Pinto e Gálvez ([1996]). Na abordagem lingüística constituída, principalmente, pelos estudos do texto e sua estrutura textual, o direcionamento realizou-se a partir de Van Dijk (1992) e Kobashi (1994).

Adotando-se a visão interacionista do processo de leitura, conforme Cavalcanti (1989) e Giasson (1993) foi realizado estudo com aborda-

gem da leitura documentária a partir da perspectiva de cada uma das três variáveis: o texto, o leitor e o contexto. A investigação, nos dois eixos de abordagem cognitiva e sociocognitiva, acompanhou a perspectiva interacionista entre as três variáveis, texto-leitor-contexto, passando a assentar, nesse tripé, toda a visão acerca da leitura documentária.

Dessa forma, examinaram-se e organizaram-se todas as condições e requisitos que se referem especificamente à leitura documentária e ao indexador como leitor profissional dentro de uma perspectiva diferente que, certamente, propiciou uma conduta teórica inovadora para a área de Tratamento Temático da Informação.

O Projeto de Pesquisa "Leitura em análise documentária: uma contribuição à formação do indexador", realizado no período de 2000 a 2004 teve a proposição de elaborar um programa de orientação à formação e capacitação do indexador em leitura documentária para alertar os responsáveis pela formação de indexadores/resumidores, ou seja, os Cursos de Graduação e de serviços de análise da informação para a necessidade de uma preparação conceitual e filosófica em torno da concepção de análise orientada para o conteúdo e para a demanda, contendo subsídios teóricos e metodológicos tanto da área de leitura quanto de indexação, observando-se os aspectos cognitivos, lingüísticos e lógicos que influem na análise de assunto. Os resultados dessa pesquisa propiciaram a elaboração do Manual de orientação ao uso do Modelo de Leitura Documentária (Apêndice 1) após aplicação com leitores profissionais proficientes e alunos de graduação aprendizes. O Manual de Orientação e o Modelo de Leitura Documentária são parte integrante do "Programa de Orientação à formação e capacitação do indexador em leitura documentária" e foram experimentados várias vezes para obter o aprimoramento necessário.

A elaboração do programa de orientação para a leitura documentária implicou a necessidade de investigação:

Primeira etapa (2000 a 2002)

- Fundamentação teórica em prática de indexação com enfâse em leitura documentária;
- Os serviços de análise e sua política de indexação;

- ♦ Formação e capacitação do indexador em leitura documentária no Brasil: bases conceituais e técnicas;
- ♦ A concepção e objetivos de leitura para o indexador: influência da formação do indexador na leitura documentária.

Segunda etapa (2002 a 2004)

- ♦ O indexador como leitor;
- ♦ Compreensão de leitura documentária em análise de assunto para identificação e seleção de conceitos;
- ♦ Formação do indexador em leitura documentária;
- ♦ Diretrizes do programa de orientação à formação do indexador em leitura documentária.

O Programa de orientação à formação do indexador a ser adotado por Cursos de biblioteconomia e cursos de educação continuada considera que a leitura documentária torna o indexador um leitor profissional uma vez que seu conhecimento prévio está vinculado ao contexto específico de seu trabalho e à sua formação profissional.

A proposição do Programa de Orientação à formação do indexador em leitura documentária elaborado com subsídios teóricos e metodológicos tanto da área de leitura quanto da indexação e observando-se os aspectos cognitivos, lingüísticos e lógicos que influem na análise de assunto, tem os seguintes objetivos:

- contribuir com a elaboração de diretrizes teóricas e metodológicas sobre leitura documentária para a orientar a formação de indexadores; e
- indicar procedimentos de elaboração de modelos metodológicos para a leitura documentária;

O Programa de Orientação à formação do indexador em leitura documentária (Apêndice 1) desenvolveu-se a partir da necessidade de uma preparação conceitual e filosófica em torno do processo de indexação e da importância de uma leitura estratégica voltada para os objetivos da análise documentária. Constitui-se de duas partes principais: a parte conceitual sobre Leitura em análise documentária e a parte metodológica dedicada à orientação para a elaboração de um modelo de leitura documentária para textos científicos.

1. Partes do Programa de Orientação

1.1 Parte conceitual: subsídios teóricos e metodológicos

A parte conceitual consta essencialmente de subsídios e recomendações à formação em leitura documentária, organizados dentro da visão interacionista das variáveis texto, leitor e contexto, contendo principalmente:

1 Leitura e compreensão de leitura

- Abordagens conceituais sobre a leitura e compreensão de leitura;
- Perspectivas interdisciplinares da leitura para fins documentários;
- Desenvolvimento de habilidades e estratégias de compreensão de leitura.

2 Leitura documentária em análise de assunto

- Leitura documentária como leitura profissional;
- Estratégias de leitura documentária para compreensão do conteúdo com objetivo da demanda.

3 Leitura em análise documentária: interação das variáveis texto, leitor e contexto

3.1 O texto: tipologias e estruturas textuais

3.2 O leitor: o indexador como leitor profissional

a) Para o indexador como leitor:

- Auto-conhecimento em leitura
- Desenvolvimento de habilidades e estratégias de leitura
- Aprofundamento de seu conhecimento prévio (lingüístico/ textual)

b) Para o indexador como leitor profissional:

- Compreensão do processo de análise documentária e da função da linguagem documentária;
- Domínio da linguagem documentária para conhecimento prévio da área de assunto;
- Indexação e análise de assunto: as concepções de análise derivadas da formação e da capacitação;
- Identificação de conceitos e tematicidade: modelos de análise conceitual;
- Seleção de conceitos durante a análise de assunto orientada para a demanda: perspectiva do usuário e da política de indexação do sistema de informação;
- Estratégias de leitura documentária em análise de assunto;
- Experimentações de leitura documentária em área especializadas;
- Conhecimento das tipologias documentárias e suas estruturas textuais;
- Análise de assunto, formação orientada para o conteúdo e para a demanda.

3.3 O contexto: físico, psicológico e sociocognitivo

a) Contexto físico: diagnóstico de infra estrutura e objetivos dos serviços de análise dos sistemas de informação;

b) Contexto psicológico: intenção, interesse e objetivos do indexador na leitura documentária para indexação;

c) Contexto sociocognitivo: o manual de indexação com as regras, procedimentos e política de indexação do sistema de informação; a linguagem documentária adotada pelo sistema.

1.2 Parte metodológica: modelo de leitura para indexação

- Identificação e seleção de conceitos: o processo de leitura documentária em Modelo de Leitura.
- Evidências da complementaridade das concepções orientadas para o conteúdo na fase de identificação de conceitos e para a demanda na fase de seleção de conceitos

A observação sobre a falta de procedimentos sistemáticos para a identificação de conceitos reforça a evidência, já mencionada a partir do estudo de aplicabilidade da análise conceitual do PRECIS, de que é necessária uma metodologia de análise de assuntos durante a leitura documentária para diminuir dificuldades. Considerando, ainda, que a exploração da estrutura textual foi uma estratégia observada freqüentemente na leitura documentária dos indexadores, indicamos a elaboração de modelo de leitura documentária em indexação de textos científicos com a possibilidade de uso combinado da exploração da estrutura textual com questionamento para identificação de conceitos.

O modelo de leitura é uma proposição de aprimoramento e evolução da metodologia de indexação a partir da metacognição do indexador que poderá ser oferecida na formação do indexador em leitura documentária. A divulgação dos resultados do estudo quanto ao processo de leitura documentária aos responsáveis pela formação e capacitação do indexador - Cursos de Graduação e Sistemas de análise da informação - visa à disseminação da importância e da influência que a leitura exerce sobre todo o desempenho da atividade de análise documentária.

Os projetos de pesquisa desenvolvidos no período de 2004 a 2010, do eixo II, tiveram abordagem sociocognitiva sob influência da perspectiva interacionista entre as três variáveis, texto-leitor-contexto.

O projeto "A leitura documentária na formação inicial do indexador: a abordagem sóciocognitiva na investigação de estratégias de ensino" (2004-2007) teve a proposta de investigação do uso de estratégias de ensino de leitura documentária para a formação inicial do indexador. Para isso, foi desenvolvida a aplicação pedagógica do modelo de leitura a partir de duas variáveis: o uso do Protocolo Verbal Interativo como recurso de aprendizagem de indexadores aprendizes e a vinculação do contexto profissional em abordagem sociocognitiva para idealização dos objetivos de leitura documentária. A leitura documentária na formação do indexador foi investigada por revisão de literatura para considerações acerca de: aspectos sócio-cognitivos e elaboração de metodologias de ensino com aplicação de Protocolo Verbal como recurso de aprendizagem.

Os estudos de perspectiva sociocognitiva forneceram indicações de estratégias de ensino que consideram o contexto como facilitador da compreensão de leitura. Para a vinculação do contexto profissional, considerou-se a interação do usuário individual com o ambiente social/organizacional. A pesquisa desenvolveu a aplicação do protocolo verbal individual e protocolo verbal em grupo com alunos do curso de Biblioteconomia.

Os resultados esclareceram as dúvidas surgidas quanto ao texto utilizado, à compreensão dos sujeitos como um todo e ao Modelo de Leitura Documentária para indexação de artigos científicos provocando seu aprimoramento e conseqüente adequação de seu Manual Explicativo como Manual de Ensino. Considerando a abordagem sócio-cognitiva, foi aplicado o protocolo verbal interativo, com o objetivo de auxiliar o indexador aprendiz, mediante a interação com o profissional experiente, esclarecendo suas dúvidas no momento da indexação, com o uso do modelo de leitura.

No projeto seguinte, "O contexto da leitura documentária de indexadores de bibliotecas universitárias em perspectiva sócio-cognitiva para a investigação de estratégias de ensino" (2007-2010), ainda no Eixo II de abordagem sociocognitiva, o enfoque foi dado ao contexto.

Nesse sentido, o contexto de leitura documentária na formação inicial do indexador precisou de fundamentação teórica e metodológica baseada no conhecimento de Ciência da Informação e áreas interdisciplinares para o desenvolvimento de estratégias de ensino que proporcionem, ao aprendiz, a visão da demanda em perspectiva sociocognitiva.

Realizou-se, com esse objetivo, a investigação do contexto de tratamento de conteúdo de bibliotecas universitárias, mediante perspectiva sociocognitiva da Análise de Domínio, com enfoque na leitura documentária para análise de assunto de livros. A finalidade dessa investigação foi a elaboração de um manual de ensino em leitura documentária com procedimentos de análise de assunto e de observação de contextos profissionais mais amplos.

O estudo do contexto de tratamento de conteúdo de bibliotecas universitárias com base em procedimentos de Análise de Domínio em Ciência da Informação pelas abordagens de indexação e recuperação em domínios específicos teve como enfoque a indexação no catálogo das bibliotecas universitárias da UNESP.

O desenvolvimento desse projeto precisou da elaboração do Projeto "Política de tratamento da informação documentária da rede de bibliotecas da UNESP", que realizou-se no período de 2006 a 2009, para a aplicação da metodologia que consistiu de estudo diagnóstico para coleta de dados com uma amostra de nove bibliotecas universitárias do sistema de bibliotecas da UNESP em três áreas do conhecimento – Odontologia, Engenharia Civil e Pedagogia.

Para realizar a coleta e análise dos dados a estrutura da investigação foi composta de três partes: funcionamento do tratamento de informações documentais na perspectiva da gerência do sistema de bibliotecas universitárias; procedimentos do tratamento de informações documentais na perspectiva do catalogador; e, avaliação do acesso e recuperação da informação on-line pelo usuário à distância.

A primeira parte constou de aplicação de questionário junto aos dirigentes; a segunda parte foi realizada com aplicação da metodologia introspectiva de Protocolo Verbal Individual junto aos catalogadores e de Protocolo Verbal em Grupo com grupos formados, em cada biblioteca, pelo dirigente,

catalogador, bibliotecário de referência, pesquisador e aluno de graduação; e a terceira parte constou da aplicação do Protocolo Verbal Individual com alunos de graduação de cada uma das três áreas do conhecimento.

Os resultados demonstraram que os catalogadores seguem uma metodologia sistematizada e consolidada pela literatura para catalogação de forma (tratamento físico), porém não demonstram metodologia para a catalogação de assunto (tratamento temático). Além disso, devido à automatização dos processos de catalogação houve uma unificação dos procedimentos técnicos da biblioteca – a seção que cuidava da forma dos documentos (catalogação) e a seção que tratava do conteúdo (classificação e indexação), o que não ocorria quando os serviços ainda eram feitos manualmente. A recuperação da informação apresenta falhas no que diz respeito à pesquisa por assunto, além disso, a linguagem documentária não atende de maneira satisfatória à demanda, não havendo compatibilidade entre os termos utilizados na catalogação de assunto e aqueles utilizados nas pesquisas pelos usuários.

Concluiu-se que a política de tratamento da informação documentária e tudo aquilo que a compõe – metodologias para representação descritiva e temática, linguagens documentárias, estudo de usuário, entre outros – necessita de maior atenção no que diz respeito à elaboração de normas, procedimentos e técnicas, sua implantação, implementação e avaliação para que a função das bibliotecas universitárias seja mantida e atualizada de acordo com as mudanças da sociedade.

Para isso, foi desenvolvida a adaptação do Modelo de Leitura Documentária para indexação na catalogação de assuntos de livros em bibliotecas universitárias e seu Manual de ensino com base na revisão de literatura e análise de procedimentos de catalogação de assuntos a partir dos protocolos verbais individuais dos catalogadores das bibliotecas universitárias e elaboração da metodologia de estudo do contexto de leitura documentária para ensino de indexação.

Os resultados do diagnóstico do contexto de tratamento temático da informação documentária coletados em nove bibliotecas da UNESP foram publicados no artigo "O contexto sociocognitivo do catalogador em bibliotecas universitárias: perspectivas para uma política

de tratamento da informação documentária" (FUJITA, RUBI, BOCCATO, 2009), e do Livro Eletrônico "Indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias" (FUJITA, 2009) publicado pela Editora da UNESP em parceria com a Pró-Reitoria de Pós-Graduação, no qual constam, além dos resultados dos relatórios dos dois projetos de pesquisa, também, os principais resultados obtidos de duas teses de doutorado e uma dissertação de mestrado. As repercussões dessa pesquisa propiciaram as condições necessárias para o desenvolvimento dos projetos de pesquisa "A política de indexação para bibliotecas" (2010-2018) e "Linguagem de indexação na perspectiva da política de indexação para bibliotecas" (2014-2017).

Referências

ALBRETCHTSEN, H. Subject analysis and indexing: from automated indexing to domain analysis. *The Indexer*, London, v. 18, n. 4, p. 219-224, oct. 1993.

ANDERSON, J. R. The architeture of cognition. Cambridge: Harvard University Press, 1983. apud PINTO, M., GÁLVEZ, C. *Análisis documental de contenido:* processamiento de información. Madrid : Síntesis, [1996].

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 12676:* Métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, 1992.

BAMBERGUER, R. *Como incentivar o hábito de leitura*. São Paulo: Ática, 1987.

BEAUGRANDE, R. *New foundations for a science of text and discourse*: cognition, communication and freedom of access to knowledge and society. Norwood: Alex, 1997.

______.; DRESSLER, W.U. *Einführung in die textlingúistik*. Tübingen: Verlag, 1981. Trad. Inglesa: Introduction to textlinguistics. apud KOCH, I.G.V. *A coesão textual.* São Paulo: Contexto, 1989.

BEGHTOL, C. Bibliographic classification theory and text linguistics: aboutness analysis, interxtuality and the cognitive act of classifying documents. *J. Doc.*, London, v. 42, n. 2, p. 84-113, 1986.

BIREME. *Manual de indexação para a base de dados LILACS*. São Paulo: BIREME, 1988.

BROWN, A. L. Metacognitive development and reading. In: SPIRO, R. et al. (Org.). *Theorical issues in reading comprehension*. New Jersey: Lauwrence Erlbaum Associates Publishers, 1980. p. 453-481.

BÜRK, K. et. al. *INIS*: manual for subject analysis. Viena: IAEA-INIS, 1995.

CAVALCANTI, M. C. *I-n-t-e-r-aç-ã-o leitor-texto*: aspectos de interpretação pragmática. Campinas: UNICAMP, 1989.

CINTRA, A. M. C. Estratégias de leitura em documentação. In: SMIT, J. W. (Coord.). *Análise documentária*: a análise da síntese. 2.ed. Brasília: IBICT, 1987. p. 29-37.

DELL, G.S; REICH,D. Storage in sentence production. A analysis of speech error data. Journal of Verbal Learning and Verbal Behavior, n. 20, p. 611-629, 1981. apud PINTO, M.; GÁLVEZ, C. *Análisis documental de contenido:* processamiento de información. Madrid : Síntesis, [1996].

FAERCH, C; KASPER, G. Process and strategies in foreing language and communication. *Interlanguage Studies Bulletin,* Utrecht, v. 5, p. 47-118, 1980 apud NARDI, M. I. A. *As expressões metafóricas na compreensão de texto escrito em língua estrangeira*. 1993. 260 f. Dissertação (Mestrado em Linguística Aplicada ao ensino de línguas)- Pontifícia Universidade Católica, São Paulo.

FARROW, J. A cognitive process model of document indexing. *J.Doc.*, London. v. 47, n. 2, p. 149-166, 1991.

______. Propositional analysis and macrorules for indexing. *Library Review*, Pittsburgh, v. 45, n. 1, p. 6-15, 1996.

______. All in the mind: concept analyis in indexing. *The Indexer*, London, v. 19, n. 4, p. 243-247, 1995.

FODOR, J.A. The modularity of mind. Cambridge, Mass.: The MIT Press, 1983. apud PINTO, M.; GÁLVEZ, C. *Análisis documental de contenido:* processamiento de información. Madrid : Síntesis, [1996].

FOSTER, K.I. Levels of processing and the structure of the language processor. In: COOPER, W.W.; WALKER, F.T. (Ed.). *Sentence Processing:* psycholinguistic studies present to Merril Garret. Hillsdale: LEA, 1979 apud PINTO, M., GÁLVEZ, C. *Análisis documental de contenido:* processamiento de información. Madrid : Síntesis, [1996].

FUJITA, M. S. L. Análise e síntese documentárias para compreensão de leitura de textos didáticos: uma proposta de aplicação do sistema de indexação Precis. *INFORMARE*: Cadernos do Programa de Pós-graduação em Ciência da Informação, Rio de Janeiro, v. 5, n. 2, 1999a.

______. *Leitura em análise documentária*. 1999b. 123 f. Relatório parcial (Projeto Integrado de Pesquisa) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista; CNPq, Marília.

FUJITA, M. S. L. *A indexação de livros*: a percepção de catalogadores e usuários de bibliotecas universitárias. São Paulo : Cultura Acadêmica, 2009, V.1.

FUJITA, M. S. L.; NARDI, M.I.A.; SANTOS, S. A leitura em análise documentária. *Transinformação,* Campinas, v. 10, n. 3, p. 13-31, set./dez. 1998.

FUJITA, M. S. L., RUBI, M. P., BOCCATO, V. R. C. O contexto sociocognitivo do catalogador em bibliotecas universitárias: perspectivas para uma política de tratamento da informação documentária. *Datagramazero* (Rio de Janeiro) abril, 2009.

GARRET, M.F. The analysis of sentence production. In: BOWER, G. (Ed.). *The psychology os learning and motivation*: advances in research and theory. Nueva York: Academic Press, 1975. v. 9 apud PINTO, M.; GÁLVEZ, C. *Análisis documental de contenido:* processamiento de información. Madrid: Síntesis, 1996.

GIASSON, J. *A compreensão na leitura*. Lisboa: Asa, 1993.

GINEZ DE LARA, M. L. *A representação documentária*: em jogo a significação. São Paulo, 1993. 133 f. Dissertação (Mestrado em Ciência da Comunicação)– Escola de Comunicação e Artes, Universidade Paulista. apud KOBASHI, N. Y. *A elaboração de informações documentárias:* em busca de uma metodologia. 1994. 195 f. Tese (Doutorado em Ciências da Comunicação)– Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo.

GOODMAN, K. S. Psycholonguistic universals in the reading process. In: PRIMSLEURS, P.; QUINN, T. (Ed.). *The psychology of second language learning*. Cambridge: Cambridge University Press, 1971. apud PINTO, M., GÁLVEZ, C. *Análisis documental de contenido:* processamiento de información. Madrid : Síntesis, [1996].

GOUGH, P. B. One second of reading. In: KAVANAGH, V. F.; MATTINGLY, I.G. (Ed.). *Language by ear and eye*: the relationships between speech and reading. [S.l.] : MIT, 1972.

GRICE, H. P. Logic and conversation. In: COLE, P.; MORGAN, P. *Syntax and semantics*: speech acts. New York: Academic Press, 1982. p. 41-58.

KATO, M. A. *O aprendizado da leitura*. São Paulo: Martins Fontes, 1985a.

______. *No mundo da escrita:* uma perspectiva psicolingüística. São Paulo: Ática, 1986.

______. Uma visão interativa da legibilidade. *Ilha do Desterro,* Florianópolis, v. 13, p. 57-66, 1985b.

KLEIMAN, A. *Texto e leitor*: aspectos cognitivos da leitura. 7.ed. Campinas: Pontes, 2000.

______. *Leitura*: ensino e pesquisa. Campinas: Fontes, 1989.

KOBASHI, N. Y. *A elaboração de informações documentárias:* em busca de uma metodologia. 1994. 195 f. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo.

KOCH, I.G.V. *A coesão textual*. São Paulo: Contexto, 1989.

______. *Desvendando os segredos do texto*. São Paulo: Cortez, 2002.

LEFFA, V. J. *Aspectos da leitura*. Porto Alegre: Sagra. 1996.

MARSLEN-WILSON, W.D. The limited compatibility of linguistic and perceptual explanations. In: GROSSMAN, R.; SAN, J.; VANCE, T. (Ed.). *Papers of parasession on functionalism*. Chicago: Chicago Linguistic Society, 1985 apud PINTO, M., GÁLVEZ, C. *Análisis documental de contenido:* processamiento de información. Madrid : Síntesis, [1996].

MASSON, M.E.J. Cognitive pocesses in skimming stories. *Journal of Experimental Psychology*: learning memory and cognition, v. 8, p. 1982. apud FARROW, J. A cognitive process model of document indexing. *J. Doc*., London, v. 47, n. 2, p.149-166, 1991.

MITCHELL, D.C. *The process of reading*: a cognitive analysis of fluent reading and learning. Nueva York: Wiley, 1982. apud PINTO, M., GÁLVEZ, C. *Análisis documental de contenido:* processamiento de información. Madrid : Síntesis, [1996].

NARDI, M. I. A. *As expressões metafóricas na compreensão de texto escrito em língua estrangeira*. 1993. 260 f. Dissertação (Mestrado em Lingüística Aplicada ao Ensino de Línguas) - Pontifícia Universidade Católica, São Paulo.

NAVES, M. M. L. Análise de assunto: concepções. *Revista de Biblioteconomia de Brasília*, Brasília, v. 20, n. 2, p. 215-226, jul./dez, 1996.

OXFORD, R. *Looking at language learning strategies*. Language learning. New York: Newbury House Publishers, 1989 apud NARDI, M. I. A. *As expressões metafóricas na compreensão de texto escrito em língua estrangeira*. 1993. 260 f. Dissertação (Mestrado em Linguística Aplicada ao ensino de línguas) - Pontifícia Universidade Católica, São Paulo.

PALINSCAR, A. S.; BROWN, A.L. Reciprocal teaching of comprehension – fostering and comprehension – monitoring activities. *Cognition and Instruction*, Hillsdale, v.2, p. 117-175, 1984.

PINTO MOLINA, M. *Análisis documental:* fundamentos y procedimientos. 2.ed.rev.aum. Madrid : EUDEMA, 1993.

PINTO, M.; GÁLVEZ, C. *Análisis documental de contenido:* processamiento de información. Madrid : Síntesis, [1996].

RUMELHART, D.E. Toward an interactive model of reading. In: DORMICI, S. (Org). *Attention and performance XL*. [S.l.]: Lawrence Esrlbaum Associates, 1977.

______.; ORTONY, A. The representation of knowledge in memory. In: ANDERSON, R.C.; SPIRO, R.J.; MONTAGUE. W.E. (Ed.). *Schooling and the acquisition of knowledge*. Hillsdale: Lawrence Erlbaum Associates, 1977.

SMITH, F. *Compreendendo a leitura*: uma análise psicolinguística da leitura e do aprender a ler. Tradução de Daise Batista. Porto Alegre: Artes Médicas, 1989. Tradução de: Understanding reading.

STEMBERGER, J.D. Na interactive activation model of language production. In: ELLIS, A.W. (Ed.). *Progress in psychology of language*. Londres: LEA, 1985. v. 3 apud PINTO, M., GÁLVEZ, C. *Análisis documental de contenido:* processamiento de información. Madrid : Síntesis, [1996].

TARDIFF, J. La evaluación del saber-ler: un asunto más de competencia que de actuación. In: RODRIGUÉZ, E; LAGER, E. (Comp.). *La lectura.* Santiago de Cali: Universidad del Valle, 1997. p. 165-194.

VAN DIJK, T. A *La ciência del texto:* um enfoque interdisciplinário. Tradução de Sibila Hunzinger. Barcelona: Paidos, 1992. Tradução de: Tekstwetenschap. Een Interdisplinaire inleiding.

______. Relevance assingnment in discourse comprehension. *Discourse Processes,* v.2, p.113-126, 1979. apud BEGHTOL, C. Bibliographic classification theory and text linguistics: aboutness analysis, interxtuality and the cognitive act of classifying documents. *J. Doc.*, London, v. 42, n. 2, p. 84-113, 1986.

VEGA, M. et al. Lectura y Comprensión: una perspectiva cognitiva. Madrid: Alianza, 1990 apud PINTO, M.; GÁLVEZ, C. *Análisis documental de contenido:* processamiento de información. Madrid : Síntesis, [1996].

Processos cognitivos na leitura documental: o que *faz* o indexador quando lê?

Ana Lúcia Terra

Introdução

Um ponto nuclear na análise de assunto é a compreensão do texto, a qual depende do processamento da informação por parte do indexador, resultante do ato leitor. Assim, o estudo da leitura realizada pelo indexador torna-se um tópico fundamental para desenhar uma metodologia que possa amparar e melhorar a compreensão do texto e, consequentemente, otimizar o processo de análise de assunto e os seus produtos. Neste pressuposto, um dos fatores mais relevantes a interferir no processo de análise de assunto é o fator humano, envolvendo aspetos cognitivos, linguísticos ou socio-culturais, na medida em que o ato leitor é realizado por indivíduos.

Se Moura (2004) qualificou o leitor-bibliotecário como uma "estranha criatura", Waller (1999, p. 35) explicitou as premissas dessa estranheza ao declarar «(…) dans sa vie professionnelle, le documentaliste ne lit plus pour lui, mais pour d'autres; il ne choisit pas ce qu'il a à lire et ne lit plus jamais un texte dans son intégralité». Este é um leitor-mediador na medida em que o resultado da sua leitura irá possibilitar aos utilizadores dos serviços de informação o acesso aos documentos através das representações de assunto (REDIGOLO, FUJITA, 2015). Assim, o indexador posiciona-se no centro de uma rede de interações, visíveis e invisíveis, pois é um mediador com uma dupla preocupação de partilha já que tem de fazer uma ligação ao autor e ao

utilizador final. Nesta perspetiva, Fondin (2006) enquadra a leitura documental no contexto da comunicação interativa e da leitura "partilhada" bem como da intersubjetividade. Na sua opinião, este último conceito é especialmente importante porque «l'intersubjectivité permet ainsi de prendre en compte à la fois le caractère "subjectif" de la lecture, c'est-à-dire l'espace de débordement reconnu à chaque lecteur du fait de sa singularité, et le souci d'une saisie "objective" du contenu du texte du fait du partage communicationnel attendu» (FONDIN, 2006, p. 27).

Portanto, coloca-se a questão do que será ler para o profissional da informação, especificamente quando exerce funções de indexação. No presente texto, iremos procurar responder a esta interrogação. Para o efeito, iremos começar por enquadrar genericamente a leitura documental no processo de indexação, nomeadamente na análise de assunto. De seguida, procuraremos sublinhar brevemente a complexidade das operações que ocorrem no cérebro do leitor. A análise dos processos cognitivos que ocorrem aquando da leitura documental, à luz da classificação que Giasson (1993) desenhou para a compreensão da leitura em ambiente de aprendizagem, será o corolário deste percurso, que terminará com algumas considerações finais.

O processo de indexação como contexto da leitura documental

Segundo Fujita (1999), os indexadores estão sujeitos a condições específicas de leitura, nomeadamente em termos de limite de tempo, de propósito definido, de produção de resultados, de vinculação a um sistema documental comprometido com as necessidades de pesquisa dos seus utilizadores, os quais irão condicionar os seus procedimentos de leitura. Neste sentido, precisam de realizar uma leitura bem estruturada e seguir um método sistematizado que apoie a identificação dos conceitos do documento, permitindo a compreensão da temática, e o reconhecimento dos termos representativos equivalentes usando a linguagem documental em uso no serviço de informação. A finalidade específica da leitura realizada pelo profissional da informação é igualmente sublinhada de forma clara por Moura (2004, p. 164) que postula uma "leitura técnica" consistindo «(…) na abordagem global dos itens informacionais, tem por objetivo recolher

os dados que permitirão o estabelecimento da representação desses itens nos sistemas de informação» e que procura «(…) a reconstituição bruta da informação veiculada no texto original». Assim, a leitura documental enquadra-se no processo de indexação, pelo que importa caracterizá-lo genericamente para melhor a compreender.

A revisão da literatura evidencia que não há unanimidade nas etapas que configuram o processo de indexação, sendo certo que independentemente da efetividade de cada uma, para os indexadores mais experientes, as etapas ocorrem frequentemente de modo simultâneo. Vários autores (SOUSA, FUJITA, 2014; MAI, 1999, 2001; entre outros) concluíram que o processo de indexação é apresentado como envolvendo duas, três ou quatro etapas. Assim, a abordagem baseada em duas etapas considera a existência de uma fase de determinação do conteúdo do documento e outra de tradução do assunto usando uma linguagem de indexação. A perspetiva que considera três etapas inclui uma etapa intermédia, a da reformulação do assunto do documento na linguagem natural. Já a corrente que preconiza quatro etapas identifica a determinação do assunto, a sua reformulação com termos da linguagem natural, a reformulação desse assunto em linguagem de indexação e, por fim, a sua tradução usando termos da linguagem de indexação.

Por seu lado, Hudon (2013, p. 34-40), preconiza que o processo de tratamento dos itens informacionais, tendo em vista a representação do seu conteúdo, compreende cinco etapas sucessivamente interdependentes. Assim, distingue as seguintes etapas: familiarização com o documento, análise, tradução do assunto em termos de indexação, validação e registo dos dados. O resultado deste procedimento deve ser uma representação concisa, completa, clara e fiável do(s) assunto(s) abordado(s) no documento. Na etapa da familiarização com o documento, o indexador identifica as principais características do recurso informacional e define o tipo de tratamento mais adequado, idealmente seguindo as orientações explicitadas numa política de indexação de âmbito institucional. Na fase da análise em sentido estrito, são determinados os assuntos do documento, são selecionados os que devem ser representados no catálogo ou no índice e decompõem-se os assuntos a serem representados em cada um dos seus elementos. Na etapa da tradução em linguagem de indexação, os assuntos selecionados são

representados através dos termos usados na(s) linguagem(ns) usadas pelo sistema de informação. A quarta etapa, da validação, visa garantir a eficácia do resultado do processo de tratamento analítico, recorrendo-se a instrumentos e métodos de verificação e aprovação. Por fim, o registo dos dados dá origem ao resultado de todo o processo, o qual pode ser uma entrada num repositório de metadados, um registo bibliográfico ou metainformação incluída no próprio recurso informacional objeto de todo o processo.

Contudo, como notou Mai (2001), entender a indexação como um conjunto simples de etapas claramente identificadas implica ignorar a complexidade deste processo. Por outro lado, sublinha que o processo de indexação não produz uma representação neutra e objetiva do(s) assunto(s) de um documento mas antes uma interpretação da representação do documento. Só entendendo este pressuposto se poderá apreender a natureza do processo de indexação. Para este autor (MAI, 1999), é fundamental compreender a índole interpretativa e incerta do processo de indexação. Neste pressuposto, criou um modelo do processo de indexação constituído por quatro elementos e três etapas (MAI, 2001). O primeiro elemento é o item informativo que origina o processo de indexação, é o documento que vai ser analisado. A primeira etapa corresponde ao processo de análise do documento e consiste no exercício de examinar o item informativo, incidindo sobre partes específicas (título, sumário, resumo, etc) de modo a identificar o assunto. O segundo elemento é um produto resultante da primeira etapa e abrange as representações mentais que o indexador formulou acerca do assunto do documento. É ainda um conjunto desordenado de impressões, frases, termos que foram surgindo durante o processo. A segunda etapa é uma resposta desencadeada no indexador pelo segundo elemento e consiste no processo de descrição do assunto. É uma tentativa de criar uma formulação coerente do assunto por meio da linguagem. Se o resultado da primeira fase é um conjunto desordenado de impressões mentais, frases e termos, o produto da segunda fase é resultante de um esforço concertado para lhes conferir ordem e estrutura. Assim, o terceiro elemento, produzido no decurso da segunda fase, consiste numa formulação do assunto do documento usando a linguagem. A terceira fase é denominada de processo de análise do assunto e consiste na tradução do terceiro elemento numa expressão formal usando linguagem de indexação apropriada. Implica,

por exemplo, converter uma frase numa notação da Classificação Decimal Universal, num cabeçalho de assunto ou em palavras-chave. Nesta fase, pressupõe-se a utilização correta das regras e convenções de cada linguagem de indexação. O quarto elemento é o produto da etapa anterior e consiste no(s) termo(s) de indexação ou na(s) notação(ões) atribuídas para representar o assunto do documento.

É neste contexto funcional que a leitura documental é efetivada com objetivos bem delimitados mas sem uma explicitação consciente de como é realizada por parte do indexador, nomeadamente ao nível dos processos e das operações cognitivas em ação. Ora, tudo se passa no cérebro.

O cérebro: onde *acontece* a leitura

Na escrita primitiva não existiam espaços a separar as palavras. Era a *scriptura continua*, a primeira forma de escrita, em que as palavras se sucediam ininterruptamente no espaço das linhas de todas as páginas, refletindo a origem oral da linguagem, pois não inserimos pausas entre cada palavra já que emitimos longas sequências de sílabas sem intervalos. Por essa razão, não se lia em silêncio mas mesmo assim a leitura era um operação lenta e cognitivamente intensa. A colocação de espaços entre as palavras escritas atenuou o esforço cognitivo para ler o texto, permitindo aos indivíduos ler rapidamente, de modo silencioso e com maior compreensão. Serve este breve enquadramento para sublinhar que esta fluência na leitura teve de ser aprendida e que exigiu profundas e complexas modificações no cérebro, tal como hoje em dia os recursos digitais e a Internet estão a modificar este órgão humano (CARR, 2012).

Assim, é compreensível que, tal como evidenciam muitos estudos científicos, os leitores experientes desenvolvem regiões especializadas do cérebro para decifrar o texto rapidamente. Nestes sujeitos, as áreas responsáveis pela representação e recolha visual, fonológica e semântica da informação importante estão ligadas para operar à velocidade da luz. Por exemplo, o córtex visual desenvolve atividade em conjuntos de neurónios dedicados a reconhecer imagens visuais de letras, padrões de letras e palavras em intervalos de milissegundos (CARR, 2012). Fica, portanto, evi-

dente que, para compreender o ato leitor, é fundamental o conhecimento e a compreensão das ações que ocorrem no cérebro aquando da sua prática.

A competência leitora convoca três capacidades neuronais (descodificar, compreender e interpretar), exigindo a intervenção sincronizada dos dois hemisférios cerebrais. Com efeito, no hemisfério esquerdo estão situadas as áreas responsáveis pela capacidade de descodificar e compreender. A descodificação implica mecanismos neuronais associados a sistemas de leitura, como a perceção da relação fonema-grafema e o conhecimento da sintaxe e morfologia. O compreender implica a procura na rede de significação de informação semelhante ou relacionada com a informação recebida. Envolve interpretação e compreensão de índole semântica e pragmática.

Como sintetiza Silveira (2013, p. 72), «(…) a relação de proximidade estreita entre estruturas da linguagem e da leitura acaba por ser um outro aspeto importante dentro dos factos que permitem ao cérebro ler competentemente. Não há competência leitora sem a existência prévia de uma linguagem capaz de representar conceitos, atribuir significados (semântica) e saber onde e como dar uso aos mesmos (pragmática), utilizando sons (fonologia), sintaxe e léxico (morfologia) corrente do contexto social em que se está inserido.». Já o hemisfério direito estará associado à compreensão geral da informação, entendida como interpretação que convoca a associação/integração da informação na rede de significação e propicia a criação de novos conceitos. É o hemisfério direito que faz a reconstrução, operação essencial ao ato leitor porque faculta a ização espacial das palavras e a atribuição do sentido global da informação. É aqui que se concretiza a criação de significado com base na recordação do conhecimento prévio sobre o contexto. Além disso, o córtex pré-frontal intervém na tomada de decisão sobre a ação leitora. Esta complexidade da decisão e da ação leitora é ainda amplificada com a intervenção do córtex visual e dos lobos parietal, temporal e frontal. Assim, fica evidente que o ato leitor exige a intervenção total do cérebro, apresentando-se do ponto de vista neurobiológico como altamente intrincado (SILVEIRA, 2013, p. 55-74). Esta complexidade tem de ser igualmente atendida no entendimento da leitura documental.

Processos cognitivos na leitura documental: proposta de tipificação

Apesar da sua relevância, como sublinha (HUDON, 2013) os processos cognitivos inerentes à determinação do assuntos dos itens informacionais, por parte do indexador, são pouco conhecidos porque têm sido pouco estudados. Isso não invalida a consciência de que a compreensão dos aspetos humanos convocados para as operações de representação informacional é um contributo valioso para a melhoria do exercício profissional (MOURA, 2004). Neste âmbito, a leitura documental apresenta uma visibilidade destacada porque é a atividade cognitiva que desencadeia todo o processo de análise de assunto, também designado pelas expressões análise temática, análise documentária ou análise de conteúdo (SOUSA, FUJITA, 2014).

De acordo com Giasson (1993), do ponto de vista do leitor, a leitura/compreensão de um texto envolve cinco tipos de processos: 1) microprocessos, 2) integração, 3) macroprocessos, 4) elaboração e 5) metacognição [Figura 1]. Os microprocessos sustentam a compreensão da informação contida numa frase. Os processos de integração servem para procurar criar ligações entre as proposições ou as frases, pois estes processos centram-se na procura da coerência entre as frases. Os macroprocessos estão orientados para a compreensão global do texto e, nesse sentido, focalizam-se na compreensão da coerência entre as frases. Os processos de elaboração permitem ao leitor ir além do texto. Os processos metacognitivos definem a compreensão e permitem ao leitor alinhar-se com o texto ou situação.

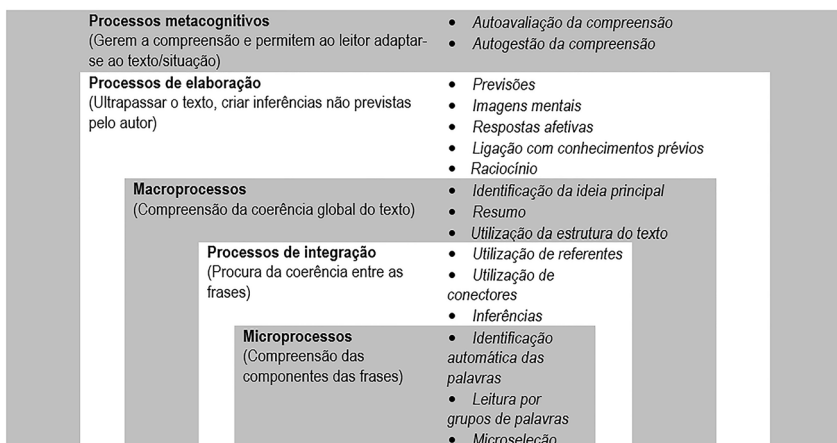



Figura 1 - Os processos de leitura documental e as suas componentes (Adaptado de GIASSON, 1993)

De seguida, iremos descrever genericamente as operações/competências inerentes a cada tipo de processo e analisar em que medida estão envolvidas ou são relevantes para a análise documental.

Os microprocessos

Os microprocessos envolvem competências ao nível da identificação das palavras, da leitura de grupos de palavras e da microseleção.

Quando a identificação das palavras é automática, a atenção do leitor pode ficar disponível para outros processos, além de tornar a leitura mais rápida. Para o trabalho de análise documental estão especialmente convocadas as duas últimas. De facto, a leitura por grupos de palavras implica a utilização de pistas sintáticas para identificar nas frases os elementos significativos de sentido, sendo que a compreensão assenta na passagem da memória de curta duração, onde são tratados os conjuntos de palavras, para a memória de longa duração, onde são armazenados. Já a microseleção consiste em decidir e memorizar a informação pertinente numa frase. Na leitura de um texto, uma parte da informação das frases lidas deve ser memorizada pela memória de curta duração, de modo a que a nova informação seja integrada na que já existe. Como o indexador tem de fazer uma leitura seletiva, já que não é viável a leitura na integra dos documentos, a leitura por grupos de palavras e a microseleção são muito relevantes. A primeira aumenta a cadência da leitura, poupando tempo e esforço cognitivo, enquanto a segunda proporciona a focalização da atenção e da memória nas componentes sintáticas e semânticas mais relevantes para extrair o(s) assunto(s) pertinentes no documento. Aliás, a literatura sobre análise de assunto sublinha que não é necessário, nem aconselhável realizar uma leitura linear, palavra por palavra, pois o leitor deve progredir no texto conforme lhe for possível antever o seguimento das ideias (FUJITA, NARDI, 1998). Estará assim, a convocar a capacidade de previsão umas das competências dos processos de integração.

Os processos de integração

Os processos de integração permitem compreender as pistas explicitas, que esclarecem a relação entre as frases (são os referentes ou anáfo-

ras e os conectores), e as relações implícitas entre as frases e as ideias,através de inferências baseadas no texto ou nas referências do leitor. As inferências permitem adivinhar o implícito e levam a uma compreensão do texto que ultrapassa o nível literal, pressupondo que o leitor recorre ao seu conhecimento pré-existente para conferir sentido a uma mensagem. Quanto mais intensa e profunda for a experiência profissional do indexador, maior será a sua capacidade de realizar inferências, podendo antecipar, ponderar ou relativizar a importância dos textos, de identificar aspetos inovadores e de selecionar a informação útil (WALLER, 1999). As inferências podem ser lógicas ou pragmáticas.

No contexto da análise documental, quanto melhor o indexador identificar as pistas explicitas mais rapidamente poderá centrar-se nas partes essenciais do texto para entender o seu sentido. Note-se ainda que a frequência da ocorrência de referentes ou anáforas ajudará a determinar o assunto do documento, sendo aliás um critério fundamental na indexação automática baseada na contagem de palavras. Contudo, este indício quantitativo não invalida a necessidade de uma aferição qualitativa do significado da ocorrência dessas palavras e dos seus equivalentes para determinar com exatidão a temática do item informativo.

No que respeita às relações implícitas, o indexador precisará de usar inferências lógicas de modo a compreender mais profundamente o sentido o texto e também irá recorrer a inferências pragmáticas, sendo que nestas últimas deve confirmar as suas suposições através de uma análise cuidada do documento objeto do processo de tratamento documental. Note-se ainda que, quanto maior for a sua experiência e o seu conhecimento das temáticas, mais inferências pragmáticas fará o indexador. Contudo, isto também aumentará a subjectividade já que este tipo de inferências está eminentemente ligado às características cognitivas e vivências de cada indivíduo.

Pinto Molina (1995, p. 229), explicitando o contributo da leitura para a elaboração de uma metodologia para elaborar resumos, também alude ao papel das inferências, considerando que «(...) these inferences allow us to create a coherent text representation, connecting meanings of different and successive sentences (...)» e acrescentado que «the inferences themselves are frequently the main point of the message». Na sua

opinião, para fins documentais, devem ser consideradas quatro categorias distintas de inferências: lógicas (usadas para definir causas e motivações ou as condições que permitem factos específicos), de avaliação (nas quais os indexadores aplicam as suas crenças às situações apresentadas), integrativas (aplicadas no momento da compreensão e baseadas em conceitos e propriedades de uma organização hierárquica) e construtivas (derivadas do conhecimento de base do indexador).

Os macroprocessos

Os macroprocessos incidem sobre a apreensão global do texto, desvendando as ligações que o tornam um todo coerente. Estes macroprocessos incluem a identificação da ideia principal, a elaboração do resumo e a apreensão/utilização da estrutura do texto, todos essenciais no trabalho de tratamento documental.

Existe uma grande diversidade no que respeita à noção de ideia principal de um texto mas podem distinguir-se duas categorias fundamentais: a) a informação é importante porque o autor a apresenta como tal, e nesse caso é uma informação textualmente importante; b) a informação é importante porque o leitor a considera como tal, tendo em conta a intenção da sua leitura, pelo que nessa situação a informação é contextualmente importante. No caso do indexador, importa que este foque a sua atenção na informação textualmente importante ainda que deva ter em conta a informação contextualmente relevante em função das necessidades dos utilizadores para os quais está a realizar o tratamento documental. Note-se ainda que pode ser relevante distinguir a temática do texto e a ideia principal do texto, a qual pode ser explicita, implícita (devendo ser inferida) ou ambígua (não estando claramente exposta).

A elaboração do resumo implica uma maior dificuldade de realização porque requer seleção e hierarquização para eliminar os elementos secundários e redundantes. O resumo consiste em reescrever um texto pré-existente, mantendo a informação importante, havendo equivalência informativa, através de uma economia de meios e a adaptação a uma nova situação de comunicação. No resumo, as operações de eliminação dos dados supérfluos, de substituição de um conjunto de elementos ou de ações

por um termo genérico, capaz de englobamento, bem como a macroselação e a invenção, escolhendo a frase que ilustra a ideia principal ou criando uma caso não exista, constituem operações estruturantes.

A estrutura do texto condiciona a organização das ideias e cada tipo de texto (narrativo, informativo, científico, etc) apresenta uma estrutura própria. O leitor precisa de apreender e utilizar essa estrutura do texto para compreender e memorizar a informação. No fundo, trata-se de formas de organização convencionadas que os leitores "armazenam" na sua mente e que ativam ao ler os textos de forma automática.

Relativamente à leitura documental, conforme sublinhado por Fujita (1999, p. 109) «Considera-se, ainda, que o leitor com facilidade de reconhecer as superstruturas textuais capta melhor as ideias principais do texto, do que um leitor que lê linearmente. Apoiando-se na estrutura textual e em seus conhecimentos prévios o leitor infere significados e levanta hipóteses que o ajudarão a apreender a temática global». Esta perspetiva é igualmente sublinhada por Naves (2001) ao enfatizar que, para realizar a leitura com fins específicos, é necessário que o indexador conheça as estruturas de diferentes tipos de textos. Para Waller (1999) esta apreensão da estrutura do texto faz-se *antes* da leitura e focaliza-se nos elementos de apresentação do texto, como os carateres tipográficos, a disposição do texto (organização dos capítulos, títulos, notas de rodapé, etc), as ilustrações (imagens, gráficos, legendas, etc), o uso de parágrafos e de alíneas, além da pontuação.

Esta identificação da estrutura dos textos permitirá categorizá-los em tipologias (monografias, artigos, textos descritivos, textos teóricos, textos problematizadores, etc) às quais será possível aplicar grelhas de leitura documental adaptadas (cf. WALLER, 1999 ou Fujita & Rubi, 2006). Estas grelhas, que constituem filtros de abordagem ao documento, incidem sobre a apresentação visual da estrutura textuale especificam as "lentes" através das quais o texto deve ser lido (ver por exemplo, FUJITA, RUBI, 2006).

Os processos de elaboração

Os processos de elaboraçãopermitem ao leitor ir além do texto, criando inferências não previstas pelo autor. Estes processos incluem cinco

componentes: previsões, imagens mentais, respostas afetivas, ligações com os conhecimentos e raciocínio

As previsões dizem respeito às ideias e incidem sobre as frases, e não sobre as palavras, podendo ser relativas ao conteúdo do texto (o que é expectável encontrar no texto) ou à sua estrutura (o que já se sabe sobre o tipo de texto em análise). As imagens mentais podem ocorrer de forma espontânea durante a leitura mas a produção consciente de imagens mentais favorece a compreensão da leitora. As imagens mentais consistem em criar na mente uma representação do que se lê. Esta atividade aumenta a capacidade de memória de trabalho durante a leitura, ao reunir os detalhes em grandes conjuntos, facilita a criação de analogias e de comparações, serve para estruturar e memorizar a informação, aumenta o grau de comprometimento com o texto e melhora o interesse pela leitura. As respostas afetivas incidem sobre as emoções convocadas pela leitura, sendo que um leitor envolvido emocionalmente com uma obra irá compreendê-la melhor. Contudo, o envolvimento afetivo pode suscitar erros levando a enfatizar aspetos que não são realmente importantes. Assim, as reações emotivas nem sempre são adequadas, dependendo do texto. A ligação com os conhecimentos significa que o leitor estabelece uma relação entre o texto e os seus conhecimentos pessoais, os quais não sendo indispensáveis à compreensão do texto, e não se afastando muito dele, podem facilitar a memorização da informação. Por fim, nos processos de elaboração estão incluídas as competências de raciocínio. De facto, compreender não é suficiente pois é necessário que os leitores formulem um julgamento crítico sobre o texto, o que implicará distinguir os factos das opiniões, conhecer as fontes de informação e estar sensibilizado para os "preconceitos" do autor.

No que respeita ao envolvimento dos processos de elaboração no decurso da análise documental, importa fazer algumas anotações. Assim, se as previsões ajudam à compreensão do texto, elas devem basear-se em pistas, motivadas por factos ou pela estrutura do texto, e nessa medida são previsões e não procedimentos de adivinhação/imaginação. Contudo, o indexador deve usar com cuidado estas pistas, nomeadamente, no que toca ao título, o qual muitas vezes, por razões de marketing editorial, entre outras, traduz uma ideia pouco precisa do conteúdo que identifica.

Por sua vez, a produção de imagens mentais implica que o leitor esteja atento ao texto e favorece a sua compreensão, apoiando a estruturação e a memorização da informação, além de facilitar a criação de analogias e de comparações, operações essenciais ao indexador não só para a apreensão do texto mas igualmente para as fases subsequentes à análise. Com efeito, a criação de analogias e de comparações serão relevantes para o momento em que os assuntos serão traduzidos em termos da linguagem documental em uso na instituição. A integração com os conhecimentos prévios do indexador é também uma operação fundamental para o procedimento de análise documental, sendo certo que a experiência em determinada área temática irá facilitar a realização de um trabalho de qualidade, apoiando a compreensão do conteúdo informativo do texto porque o leitor já possui referentes que o ajudam a enquadrar-se no assunto (REDIGOLO, FUJITA, 2015). Assim, é mais fácil compreender um texto produzido no âmbito da mesma matriz cultural ou numa mesma área científica já que os referentes são os mesmos. Com efeito, quanto mais informação um sujeito tiver sobre uma temática mais apto estará para compreender, memorizar e produzir inferências a partir da sua leitura. Ainda assim, também é verdade que os conhecimentos prévios podem afastar o leitor dos aspetos essenciais do texto, desviando o objetivo da leitura. Por outro lado, é também essencial ter cautela com os conhecimentos erróneos pois a natureza do conhecimento prévio afeta a aquisição de novos conhecimentos. Neste sentido, se o indexador tiver conhecimentos pouco precisos ou errados sobre uma matéria terá tendência a distorcer o que ler, de modo a adequar esse conteúdo aos seus quadros mentais prévios, mesmo que estejam errados. Nesta perspetiva, será importante que o indexador tenha competências para aplicar estratégias que o conduzam a comparar os seus conhecimentos com os novos dados fornecidos pelo texto. Aqui, o indexador estará a acionar a memória, comparando o que conhece com o desconhecido, tal como sublinhado por Pinto Molina (1992 apud PINTO MOLINA, 1995).

O raciocínio é igualmente uma operação nuclear no processo de análise documental, sustentando a distinção entre factos e opiniões ou identificando os aspetos denotativos da linguagem (apreensão do sentido literal) e os aspetos conotativos da linguagem (apreensão do sentido figurado da linguagem). Pinto Molina (1992 apud PINTO MOLINA, 1995)

também realça o contributo desta componente e das atividades que lhe estão associadas (indução e dedução, análise e síntese). Já as respostas afetivas, enquanto reações emocionais que fazem sobrepor a individualidade e as preferências do indivíduo, enfatizando, portanto, a subjetividade, devem ser evitadas pelo indexador no momento da sua apropriação do texto para efeitos de análise documental.

Os processos metacognitivos

Por fim, os processos metacognitivos gerem a compreensão e permitem ao leitor adaptar-se ao texto e à situação/contexto em que decorre a leitura. Genericamente, por cognição, pode entender-se «a forma como o cérebro percebe, aprende, recorda e pensa sobre toda informação captada através dos cinco sentidos» (Wikipedia). A cognição envolve a compreensão, a memorização e o tratamento da informação. Já a metacognição diz respeito ao conhecimento que o indivíduo tem sobre o seu próprio funcionamento cognitivo e as estratégias para controlar esse processo.

As competências metacognitivas desdobram-se em duas vertentes: na autoavaliação e na autorregulação. Pela autoavaliação o leitor torna-se ciente das suas capacidades e das limitações cognitivas, dos seus interesses, motivações, bem como das exigências da leitura e de estratégias adequadas para ultrapassar problemas de leitura. Trata-se, portanto, de identificar perdas da compreensão. Com a autorregulação o leitor consegue resolver a perda de compreensão pois identifica quando compreende e quando não compreende, sabe o que compreende ou não e aquilo de que necessita para compreender e sabe que pode seguir outras estratégias para compreender. Assim, o leitor consegue aferir se compreende bem, posicionando-se numa perspetiva de autogestão da compreensão, identificando e resolvendo as lacunas de compreensão.

Na opinião de Fujita (1999, 108) «(…) a leitura para fins documentários exige que o leitor-indexador seja metacognitivo por ser necessário usar estratégias de interação com o texto de forma a propiciar uma compreensão suficiente para a sua representação». Mais recentemente, Redigolo e Fujita (2015) trabalharam o levantamento de estratégias metacognitivas através a aplicação do protocolo verbal a profissionais da

informação dedicados à análise de assunto em monografias em bibliotecas universitárias. Neste contexto, concluíram que «as estratégias metacognitivas demonstram que o leitor tem consciência dos objectivos da leitura e utiliza-se de alguns artifícios para realiza-la da melhor maneira, tentando superar suas dificuldades e explorando conhecimentos» (REDIGOLO; FUJITA, 2015, p. 365).

Considerações finais

Ainda que o modelo de leitura de Giasson (Leitor-Texto-Contexto) não tenha sido desenvolvido especificamente para explicar os processos de leitura documental,tem guiado alguns estudos nesta matéria (cf. FUJITA, NARDI,1998 entre outros). Contudo, até agora, não parece ter sido aplicada a sua tipificação dos processos cognitivos que ocorrem durante o ato leitor no contexto particular da leitura documental.Neste trabalho, expusemos como a aplicação dessa tipificação dos processos cognitivos e das suas componentes/operações pode ser bastante proveitosa em termos de compreensão dos processos cognitivos apresentados pelo indexador. Esta linha de análise precisa ainda de ser aprofundada em termos teóricos e com aplicações de índole prática mas a identificação dos processos feita por Giasson parece adequada para compreender os processo cognitivos que ocorrem durante a leitura documental, apesar das suas particularidades. Na verdade, a leitura realizada com intuito de indexação não apresenta peculiaridades significativas em relação à leitura produzida em condições naturais. Assim, com base num estudo empírico, Neves, Dias, Pinheiro (2006, p. 150) constataram que «com exceção de maior uso da estratégia de resumo e de alguns sinais de maior monitoramento da compreensão, a leitura do indexador não se diferenciou da leitura de outros leitores proficientes». Os mesmos autores adiantam que os indexadores da sua amostra não evidenciam estratégias metacognitivas específicas, não praticando, portanto, uma leitura diferenciada.

Por outro lado, se considerarmos que as etapas do processo de indexação são desenvolvidas pelos sujeitos em simultâneo, sobretudo quando já são profissionais experientes, poderá ser interessante perspetivá-las não enquanto etapas ou fases mas antes como operações cognitivas. Neste sen-

tido, o indexador deve investir na consciencialização dessas operações, base para a sua análise e subsequentemente para a sua melhoria. Neste sentido, a leitura documental também pode ser encarada como um conjunto de processos que decorrem em simultâneo e não como um conjunto de fases sequenciais hierarquizadas.

Referências

CARR, N. *Os superficiais:* o que a internet está a fazer aos nossos cérebros. Lisboa: Gradiva, 2012.

FONDIN, H. La problématique de la lecture documentaire du texte informationnel. *Documentation et Bibliothèques*, *Jan-Mar*, 17–28, 2006.

FUJITA, M. S. L. A leitura do indexador: estudo de observação. *Perspectivas em Ciência da Informação*, v. 4, n. 1, p. 101-116, 1999.

FUJITA, M. S. L.; RUBI, M. P. Um modelo de leitura documentária para a indexação de artigos científicos: princípios de elaboração e uso para a formação de indexadores. *DataGramaZero- Revista de Ciência da Informação*, v. 7, n. 3, 2006.

FUJITA, M. S. L.; NARDI, M. I. A. A leitura em análise documentária. *Transinformação*, v. 10, n.3, p. 13-31, 1998.

GIASSON, J. *A compreensão na leitura.* Lisboa: Asa, 1993.

HUDON, M. *Analyse et représentation documentaires:* introduction à l'indexation, à la classification et à la condensation des documents. Québec: Presses de l'Université du Québec, 2013.

MAI, J.-E. Deconstructing the indexing process. *Advances in Librarianship*, v. 23, p. 269–298, 1999.

MAI, J.-E. Semiotics and indexing: an analysis of the subject indexing process. *Journal of Documentation*, v. 57, n. 5, p. 591–622, 2001.

MOURA, M. A. Leitor-bibliotecário: interpretação, memória e as contradições da subjetividade. *Perspectivas Em Ciência Da Informação*, v. 9, n. 2, p. 158–169, 2004.

NAVES, M. M. L. Estudo de fatores interferentes no processo de análise de assunto. *Perspectivas Em Ciência Da Informação*, v. 6, n. 2, p. 189–203, 2001.

NEVES, D. A. B.; DIAS, E. W.; PINHEIRO, Â. M. V. Uso de estratégias metacognitivas na leitura do indexador. *Ciência Da Informação*, v. 35, n. 3, p. 141–152, 2006.

PINTO MOLINA, M. Documentary abstracting: toward a methodological model. *Journal of the American Society for Information Science*, v. 46, n. 3, p. 225–234, 1995.

REDIGOLO, F. M.; FUJITA, M. S. L. A leitura profissional do catalogador e seu papel como mediadora da informação. *Informação & Informação*, v. 20, n. 3, p. 356–376, 2015.

SILVEIRA, T. *Cérebro e leitura: fundamentos neurocognitivos para a compreensão do comportamento leitor no processo educativo*. Bloco Editora, 2013.

SOUSA, B. P.; FUJITA, M. S. L. Análise de assunto no processo de indexação: um percurso entre teoria e norma. *Informação E Sociedade: Estudos*, v. 24, n. 1, p. 19–34, 2014.

WALLER, S. *L'analyse documentaire:* une approche methodologique. Paris: ADBS Éditions, 1999.

Estratégias de Leitura Documentária para Indexação: um estudo de caso com o Centro de Informações Nucleares (CIN)

Milena Polsinelli Rubi

1 Introdução

A leitura documentária, realizada pelo indexador na fase de análise, visa propiciar a identificação de conceitos para posterior representação nos produtos documentários – número de classificação, resumos e índices – que satisfaçam a demanda dos usuários.

Para isso, o leitor documentalista, apesar de não ser o especialista do assunto, interage com o texto mediante o domínio de uma linguagem documentária, da estrutura textual e da intenção do sistema de informação para leitura.

Desenvolvido dentro do Projeto Integrado "Leitura em Análise Documentária[1]", esse estudo foi realizado com o Centro de Informações Nucleares da Comissão Nacional de Energia nuclear (CIN/CNEN) que é, desde 1970, responsável pela coleta e disseminação de informações na área de energia nuclear no Brasil.

---

[1] Desenvolvido pela Dra. Mariângela Spotti Lopes Fujita, o referido projeto visou, principalmente, à obtenção de subsídios sobre os serviços de análise documentária em Centros de Informações especializados no Brasil. A primeira parte do projeto realizou estudo de caso junto à Sub-Rede Nacional de Informações em Ciências da Saúde Oral do Sistema BIREME/ OPAS para a produção da base de dados LILACS.

Durante a realização da pesquisa, o CIN apresentava dez indexadores especialistas em áreas de assunto relacionadas à Energia Nuclear (Biologia, Energia, Energia Nuclear, Engenharia Elétrica, Engenharia Química, Física, Geologia, Medicina Nuclear, Meio Ambiente e Química e outros) e em procedimentos de indexação sistematizados para a inserção de registros na base de dados do *International Nuclear Information System* (INIS), do qual o Brasil é país-membro cooperante através do CIN. Desses indexadores apenas um era é funcionário do CIN, os outros nove eram indexadores *freelancers*, tendo seus serviços de indexação terceirizados. Nenhum era bibliotecário.

Foram coletados e analisados os dados referentes ao CIN e seus indexadores. Dessa forma, e com o conhecimento teórico da atividade de análise documentária, foi possível investigar, no contexto de um centro referencial em Energia Nuclear, a prática da atividade de indexação e caracterizar o profissional que a executa, com os seguintes objetivos:

- Identificar os procedimentos de leitura documentária utilizados para exame de documento e identificação de conceitos;
- Examinar as estratégias de leitura usadas pelos indexadores para a identificação de conceitos que melhor representem o documento.

Como metodologia, foram utilizados o manual de indexação do INIS e a técnica de coleta de dados introspectivos denominada "Protocolo Verbal" ou "Pensar Alto" nos moldes de Ericsson e Simon (1987), seguida da entrevista retrospectiva. Esta técnica permite a observação das estratégias utilizadas pelos indexadores, pois privilegia o Pensar Alto durante a realização da tarefa.

2 Estratégias de leitura documentária para indexação

A leitura documentária constitui-se a atividade principal da Análise Documentária. Na indexação, o resultado será a identificação e seleção dos termos que irão representar o documento para o usuário.

O termo indexação (*indexing*) remonta à corrente inglesa e, de acordo com os "Princípios de indexação" do *World Information System for Science and Technology*[2] (UNISIST, 1981, p. 84) é"[...] a ação de descrever e identificar um documento de acordo com seu assunto."

Essa publicação do UNISIST originou a primeira norma a esse respeito publicada em 1985 pela *International Standardization for Organization* (ISO), sob número 5963, com o título "*Documentation - methods for examining documents, determining their subjects, and selecting indexing terms*".

Em 1992, a Associação Brasileira de Normas Técnicas (ABNT) publicou a tradução dessa mesma norma, sob número 12676, intitulada "Métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação". A indexação é definida pela Norma 12676 como "Ato de identificar e descrever o conteúdo de um documento com termos representativos dos seus assuntos e que constituem uma linguagem de indexação." (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992, p. 2).

Para Chaumier (1988, p. 63) "[...] a indexação é a parte mais importante da análise documentária. Consequentemente é ela que condiciona o valor de um sistema documentário." Ainda segundo o autor, uma indexação inadequada ou uma indexação insuficiente representa 90% das causas essenciais para a aparição de "ruídos" (os documentos não pertinentes à questão que são recuperados em uma pesquisa bibliográfica) ou de "silêncios" (os documentos pertinentes à questão existentes no acervo que não são recuperados).

O termo indexação é definido por Van Slype (1991) como a operação que consiste em enumerar os conceitos sobre os quais trata um documento e representá-los por meio de uma linguagem combinatória – lista de descritores livres, lista de autoridades e o thesaurus de descritores – tendo como finalidade a busca documental que será realizada a partir dos índices ou dos catálogos. Nessa definição, referido autor destaca a representação dos conceitos, por meio de uma linguagem específica com vistas ao

[2] Sistema internacional vinculado à Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) e conhecido pela sigla UNISIST.

processo de recuperação da informação por meio de índices ou catálogos, colaborando com nossa intenção de pesquisa ao estudarmos a política de indexação para construção de catálogos.

Lancaster (2004, p. 1) explica que "[...] os processos de indexação identificam o assunto que trata o documento [...]" e eles implicam "[...] a preparação de uma *representação* do conteúdo temático dos documentos." (LANCASTER, 2004, p. 6, grifo do autor).

Consideramos, portanto, que a indexação diz respeito à identificação do conteúdo do documento, por meio do processo de análise, e à representação desse conteúdo através de conceitos. Esses conceitos, por sua vez, serão representados ou traduzidos em termos advindos de uma linguagem documentária, com vistas à intermediação entre o documento e o usuário no momento da recuperação da informação, seja em índices, catálogos ou bases de dados.

O processo de indexação é composto por diferentes etapas, cujo número varia de acordo com os autores, como veremos a seguir.

O UNISIST (1981) afirma que o processo de indexação compreende dois estágios:

- o estabelecimento de conceitos tratados num documento: envolve a compreensão do conteúdo do documento como um todo; a identificação dos conceitos representativos desse conteúdo; a seleção dos conceitos válidos para recuperação.
- a tradução dos conceitos nos termos da linguagem de indexação;

Segundo Chaumier (1986) existem também duas etapas:

- o reconhecimento e extração de conceitos informativos composto pela apreensão do conteúdo do documento por meio da leitura; pela identificação dos conceitos tendo em vista os objetivos do sistema de informação as necessidades dos usuários e pela seleção de conceitos segundo a exaustividade e especificidade do sistema de informação;
- a tradução desses conceitos em linguagem natural.

Para Van Slype (1991) a indexação comporta quatro operações distintas, a saber:

- conhecimento do conteúdo do documento;
- escolha dos conceitos a serem representados, baseando-se na aplicação da regra da seletividade e exaustividade;
- tradução dos conceitos selecionados da forma em que aparecem impressos no documento, para os descritores do tesauro aplicando a regra da especificidade;
- incorporação dos elementos sintáticos.

Assim como para o UNISIST (1981) e para Chaumier (1986), para Lancaster (2004) a indexação também envolve duas etapas:

- a análise conceitual:
- a tradução.

A Norma ABNT 12676 descreve três estágios da indexação:

- exame do documento e estabelecimento do assunto de seu conteúdo;
- identificação dos conceitos presentes no assunto;
- tradução desses conceitos nos termos de uma linguagem de indexação.

Os autores divergem quanto ao número de etapas da indexação. Porém, elas tratam basicamente das mesmas operações:

- análise: leitura e segmentação do texto para identificação e seleção de conceitos;
- síntese: construção do texto documentário com os conceitos selecionados;
- representação: por meio de linguagens documentárias.

Observamos, portanto, que a etapa inicial da indexação é a análise. Essa análise é realizada por meio da leitura feita pelo indexador que procura compreender de maneira geral o documento para identificar e selecionar os termos que representarão o documento para a recuperação.

Concordando com os autores Fujita (1998; 2001; 2003) e Dias e Naves (2007), acreditamos que a leitura documentária é a etapa primordial da indexação, pois corresponde à primeira fase de abordagem do leitor indexador e desencadeará o processo de análise de assunto do texto para identificação e seleção de conceitos.

No entanto, na prática profissional do indexador, a leitura total de um documento é impraticável, já que a leitura documentária está sujeita a condições específicas: limite de tempo, propósito definido, geração de produtos e serviços e grande volume de documentos de uma mesma área de assunto.

Sendo assim, o indexador deve lançar mão de seu conhecimento prévio – linguagem documentária utilizada pelo sistema de informação, estrutura textual do documentoe assunto – e acionar estratégias durante a leitura documentária a fim de que seu objetivo seja atingido: identificação e seleção de conceitos de um documento.

As estratégias de leitura, as ações que o leitor realiza no ato de ler, têm sido definidas por vários autores. Essas estratégias segundo Faerch e Kasper (1980, citados por NARDI, 1993) são planos potencialmente conscientes do leitor para resolver algo que se apresenta como um problema na compreensão.

Para Brown (1980) estratégia é qualquer controle deliberado e planejado das atividades que levam a compreensão. A autora ainda acredita que as ações são intencionalmente selecionadas.

Kato (1987) distingue dois tipos de estratégias que definem o comportamento do leitor: as *estratégias cognitivas* que são aquelas que automáticas e subconscientes utilizadas durante a leitura fluida e as *estratégias metacognitivas* que são conscientes do leitor frente a um problema. As estratégias cognitivas de Kato (1987) são denominadas por Cavalcanti (1978) como estratégias automáticas e as metacognitivas são chamadas estratégias controladas.

Sob o ponto de vista de Nardi (1993), Brown (1980) considera estratégia cognitiva como "*skill*" (uma estratégia que teria sido adquirida em algum momento da aprendizagem e se tornado automática) e estratégia metacognitiva como estratégiasimplesmente (o uso consciente da uma "*skill*" ou de uma nova tática, em momentos de solução de problemas).

Na nossa concepção, estratégias são ações empreendidas pelo leitor no ato de ler, potencialmente conscientes, geralmente direcionadas para a solução de um problema.

Brown (1980) lista atividadesconsideradas de natureza metacognitiva:

- explicitação dos objetivos da leitura;
- identificação de aspectos importantes da mensagem;
- alocamento de atenção a áreas importantes;
- monitoração do comportamento para ver se está ocorrendo compreensão;
- engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido;
- tomada de ações corretivas quando são detectadas falhas na compreensão;
- recobramento de atenção quando a mente se distrai ou faz digressões.

Para Kato (1987) existem apenas duas estratégias metacognitivas básicas, sendo as demais apenas subtipos dessas estratégias. São elas:

- estabelecimento de um objetivo explícito para leitura;
- monitoração da compreensão tendo em vista esse objetivo.

Em se tratando de leitura documentária, consideramos importante abordar as noções de estratégias de leitura em documentação apresentadas por Cintra (1989), em que são apontados vários fatores que con-

correm para a qualidade de um texto: a manutenção do tema, a correção gramatical, a adequação lexical, a estrutura do texto.

Segundo a referida autora, na leitura para fins documentários é preciso que haja uma cooperação autor/leitor, já que o autor não prevê quem poderá ler o que ele publicou.

Sabe-se também que na leitura para fins documentários e tratamento da informação (leitura técnica), não é necessário, nem aconselhável uma leitura linear, letra por letra, palavra por palavra, o leitor avança no textona medida que consegue predizer o que vem a seguir.

O leitor deve buscar detectar a estrutura do texto. O leitor, que tem facilidade de reconhecer as superestruturas textuais, capta melhor as ideias principais do texto. Diferentemente, um leitor que lê linearmente, faz esforços desnecessários para compreender trechos isolados, mas sem apoiar-se na estrutura textual e seus conhecimentos prévios para inferir significados, e levantar hipóteses que o ajudarão a apreender a temática global. O apoio na estrutura textual permite ao leitor ser seletivo e alocar atenção a trechos importantes.

As estratégias ascendentes (*bottom-up*) e descendentes (*top-down*) citadas por Cintra (1989) mostra que o leitor, durante o processo deleitura, pode realizar dois movimentos: movimento *botton-up*, em que o leitor vai lendo na dependência do contexto escrito, ou seja, vai extraindo, linearmente, dos símbolos impressos o significado, caminhando das partes para o todo, e movimento *top-down*, no qual há maior dependência de conhecimento prévio do leitor, pois ele vai fazendo generalizações e predições a partir de "esquemas" que tem armazenados em sua memória, formulando hipóteses que ajudarão na compreensão do texto.

Kato (1987) considera leitor experiente aquele que utiliza os dois tipos de estratégias, as ascendentes (dependentes do texto, da análise cuidadosa do *input* visual) eas descendentes (baseadas no conhecimento prévio do leitor e na sua capacidade de inferência, de predição), relacionando ora um tipo, ora outro, de maneira consciente, no momento em que cada uma delas se fizer necessária. Há momentos na leitura em que um trecho difícil para o leitor exige que ele leia linear e cuidadosamente, e há outros em que apenas inferências pelo contexto permitem a compreensão sem problemas.

Além disso, acredita-se que o leitor estratégico proficiente seja aquele que além de utilizar apropriadamente estratégias ascendentes e as descendentes, mantém em mente o objetivo da leitura.

Para essa verificação, foi realizado diagnóstico e caracterização do indexador do Centro de Informações Nucleares (CIN) por meio de suas atividades de análise e leitura documentária para indexação.

3 Metodologia

3.1 Centro de Informações Nucleares (CIN)

O Centro de Informações Nucleares (CIN), departamento da Comissão de Nacional de Energia Nuclear (CNEN), faz do Brasil um país membro cooperante do *International Nuclear Information System* (INIS) que é *coordenado pela InternationalAtomic Energy Agency* (IAEA). O principal produto do desse sistema de informação é a base de dados bibliográfica INIS que, atualmente possui mais de 3 milhões de registros sendo a fonte mundial mais completa sobre literatura técnico-científica na área nuclear. (BRASIL, 2015).

O CIN "[...] é o ponto de contato nacional para encontrar informações específicas sobre a área nuclear através do acesso a serviços automatizados e à base de dados bibliográficas." (BRASIL, 2015).

A opção por um estudo de caso com esse sistema de informação partiu do pressuposto de que as atividades de análise documentária realizadas para geração da Base de Dados INIS contaria com recursos humanos treinados e com metodologia padrão.

Com a permissão e o interesse do Sr. Luiz Fernando Passos de Macedo, então gerente do Centro de Informações Nucleares, foi possível a coleta de informações necessárias à caracterização funcional desse sistema.

O funcionamento do INIS depende de seus membros cooperantes que se encarregam de:selecionar a literatura nuclear gerada em seu país;preparar os dados de entrada de acordo com as regras do INIS;enviar uma cópia do texto completo de artigos da literatura que não está disponí-

vel através dos canais comerciais convencionais;disponibilizar os serviços e produtos de informação do INIS aos usuários e/ou membros do Sistema.

Para isso, a atividade de análise documentária, compreendida pelas fases de análise, síntese e representação, é realizada no CIN para a alimentação da Base de Dados INIS. Esses registros representam toda a literatura (teses, artigos, livros, capítulos de livros, trabalhos de congressos etc.) gerada e publicada por pesquisadores brasileiros.

Para a coleta desses trabalhos são feitos contatos com as instituições da pesquisa, universidades, bibliotecas da área nuclear e até mesmo com os próprios autores para que sejam enviados ao CIN os trabalhos publicados na área de energia nuclear e afins.

Os documentos que chegam ao CIN passam por um processo de seleção de assunto para garantir que eles estejam dentro do escopo da Base de Dados INIS.

Para manter a uniformidade dos arquivos informacionais foram elaboradosmanuais de referência que contém regras, padrões, formatos e diretrizes para catalogadores e indexadores nos quais o Sistema está baseado. Conhecidos como INIS Reference Series, esses manuais estavam disponíveis em CD-ROM e, atualmente estão disponíveis no site www.iaea.org.

Os indexadores do CIN realizam o processamento temático dos documentos utilizando o INIS: Manual for SubjectAnalysis (IAEA-INIS-12) produzido em conjunto pelo INIS e ETDE (Energy Technology Data Exchange) e publicado em 1995 para substituir o INIS: Manual for Indexing (IAEA-INIS-12, Rev.2), além do INIS Thesaurus.

O Manual descreve em sequência todos os passos da análise de assunto em6 capítulos: 1 Introdução à Análise de Assunto; 2 Análise Preparatória; 3 Classificação de Assunto; 4 Resumo e Aumento de título; 5 Indexação de Assunto; 6 Diretrizes e Exemplos para Análise Temática em campos de Assuntos Particulares.

Nos procedimentos sugeridos para indexação existe apenas um item destinado à leitura do documento. Sendo assim houve a necessidade de leitura de todo o Manual para reconhecimento e síntese de aspectos relacionados à leitura.

Após esse reconhecimento, os aspectos sintetizados foram ordenados para garantir que a sequência "antes da leitura - durante a leitura - após a leitura" fosse mantida.

Recomendações do sistema para identificação de conceitos

1. O entendimento do conteúdo técnico-científico do documento é o pré-requisito básico para indexação, isto significa que o indexador deverá ser especialista na área do assunto da publicação a ser indexada, e ao mesmo tempo, familiarizado com o Thesaurus e com as regras de indexação.
2. O critério a ser sempre utilizado para interpretação do documento deverá ser inteiramente baseado no usuário.
3. O indexador deverá estar atento para identificar cada parte da literatura sob o ponto de vista do usuário
4. Identificar os conceitos principais que geralmente constituem a razão essencial da publicação do material.
5. Identificar uma, duas ou talvez três ideias principais que constituem a razão essencial da publicação do material. Cada uma dessas ideias deverá ser representada por descritores coordenados.

Leitura e identificação de conceitos

1. Leia cuidadosamente o título e resumo.
2. O mais comum é começar a indexação pelos conceitos contidos no resumo e no título e depois esquadrinhar o texto todo para encontrar itens informacionais que estão faltando no resumo ou que requerem mais precisão.
3. Esquadrinhe o texto todo para encontrar conceitos mais detalhados, prestando cuidadosa atenção ao Índice de Assunto, Introdução, Resumo, Conclusões, Gráficos e Tabelas.

4. O indexador deve decidir se a publicação consiste de partes de procedimentos com tópicos diferentes suficientemente para que uma **subdivisão** seja feita no nível bibliográfico e cada parte seja tratada como um documento separado.
5. Para encontrar o descritor mais específico apropriado, o indexador deverá checar *wordblock* do descritor considerado para ver se algum dos termos menores associados a ele são apropriados.
6. Sugere-se que uma lista de trabalho dos itens informacionais seja feita primeiramente. Durante a leitura do título e resumo e esquadrinhamento do texto todo, o indexador lista aqueles itens que representam a informação significativa. Feita esta lista, ela deverá ser "traduzida" em descritores.

Representação dos conceitos por descritores do Thesaurus

1. Cada um dos conceitos identificados deve ser representado por descritores retirados do Thesaurus
2. Selecione somente aqueles descritores que representem conceitos **realmente discutidos em detalhe** no documento e conceitos sobre os quais alguma informação útil pode ser obtida pela leitura do documento.

3.2 Protocolo verbal

Para a observação das estratégias de leitura de indexadores do CIN foi aplicada, como método de observação de processos, a técnica de coleta de dados introspectivos "Protocolo Verbal" ou "Pensar Alto" nos moldes de Ericsson e Simon (1987). Segundo os autores, as observações do processo são como observações que fornecem informações sobre passos de processamento individual, tais como verbalizações espontâneas, sequência de movimentos com os olhos, exteriorizando seus processos mentais, mantendo a sequência das informações processadas.

Cohen (1987) cita o "Pensar Alto" dentre três tipos básicos de dados provenientes da técnicas introspectivas:

- Auto-Relato: afirmações generalizadas sobre o comportamento durante a leitura;
- Auto-Observação: inspeções de comportamentos específicos de leitura, enquanto a informação ainda está sob o foco de atenção;
- Auto-Revelação: "Pensar alto", o pensamento é direta e automaticamente externalizado.

Para Ericsson e Simon (1987), o leitor pode exteriorizar seus processos mentais enquanto a informação processada estásob o foco de sua atenção. O indivíduo lê e interpreta ao mesmo tempo, exteriorizando em voz alta tudo o que "passa pela sua cabeça" durante a leitura.

Nardi (1993) esclarece que, na procura de métodos para acesso direto aos processos mentais dos indivíduos, os pesquisadores encontraram apoio na Linguística, Sociologia e Psicologia. Essas áreas utilizam métodos introspectivos(Relatos, Protocolos, Declarações Verbais), ou seja, declarações dos próprios informantes sobre como organizam e processam a informação.

Segundo Cavalcanti e Zanotto (1994), protocolos são geralmente definidos como relatos verbais dos processos mentais conscientes do informante. Em outras palavras, eles se referem ao "pensar alto" do informante enquanto realiza uma tarefa de qualquer natureza.

Os protocolos verbais foram introduzidos na pesquisa qualitativa em Psicologia a partir de 1980 e desde então sua validade como reveladores de processos mentais tem sido questionada. Durante a era do behaviorismo, eles foram banidos da ciência. Segundo Ericsson e Simon (1987), quando o cognitivismo entrou em evidência como um novo paradigma, eles ressurgiram como principal fonte de dados para a pesquisa cognitiva. Este renascimento ocorreu dentro do arcabouço teórico do processamento da informação relacionado a estudos de solução de problemas.

Eles ultrapassaram os limites da Psicologia Cognitiva e passaram a ser usados na Linguística Aplicada, área em que marcaram presença forte na pesquisa sobre leitura em língua estrangeira.

Segundo Ericsson e Simon (1987) alguns autores criticam a técnica do "pensar alto", por acreditarem que ela pode modificar os processos mentais, fazendo com que as informações dos sujeitos não sejam precisas, completas, não sejam confiáveis. Em defesa do Relato Verbal, Ericsson e Simon (1987) apoiam-se na teoria do processamento da informação e explicam que a informação é armazenada na memória a curto prazo (STM) - duração intermediária de retenção e acesso rápido à informação e na memória de longo prazo (LTM) - armazenamento e duração relativamente permanente, lenta a recuperação da informação.

O modelo proposto por Ericsson e Simon (1987) prevê que a informação recém-apreendida pelo processador central é mantida na memória de curto prazo por algum tempo e é diretamente acessível para processamento subsequente, enquanto que na memória de longo prazo a informação precisa ser recuperada antes de ser relatada. Dessa forma, as informações coletadas em Relatos Verbais são as recém-apreendidas, diretamente acessíveis para processamento subsequente.

Cavalcanti e Zanotto (1994) argumentam que a questão da confiabilidade dos dados provenientes de técnicas introspectivas é um problema de outras técnicas também. Apesar de ainda controvertida, essa técnica é, segundo Nardi (1993), o único instrumento de coleta, no momento disponível, que possibilita observar processos do leitor durante a compreensão de um texto, o que justifica a nossa escolha.

A seguir, descrevemos a metodologia do Protocolo Verbal para observação da leitura documentária.

1. Seleção do texto

A seleção do Texto-Base foi solicitada ao CIN com a recomendação de que fosse um texto ainda não indexado e que não fosse de uma

área muito específica, podendo ser indexado por qualquer profissional independentemente de sua área de atuação.

O texto selecionado pelo CIN é da área de Química publicado como artigo no periódico *Química Nova:*

FIGUEIRA, R.C.L.; CUNHA, I.I.L. A contaminação dos oceanos por radionuclídeos antropogênicos. *Química Nova*, São Paulo, v.21, n.1, p.73-77, 1998.

O texto apresenta uma figura e três gráficos além da seguinte estrutura textual:

- Título;
- Resumo (somente em inglês);
- Palavras-chaves (omitidas durante a entrevista)
- Introdução
- Dois subtítulos:
  - Comportamento dos radionuclídeos no meio marinho;
  - Níveis de radionuclídeos nos oceanos;
- Conclusão;
- Referências bibliográficas.

2. Seleção dos sujeitos

A seleção dos sujeitos indexadores para o propósito da pesquisa teve como critérios o tempo de permanência da pesquisadora no Rio de Janeiro e a disponibilidade dos indexadores externos.

De acordo com a metodologia de observação do Protocolo Verbal, a entrevista realizada com 4 indexadores constou da leitura do texto-base. Para esta leitura, a entrevistadora solicitou que a fizessem naturalmente, conforme sua preferência e rotina diária, tendo como objetivo a identificação e seleção de conceitos para a indexação, uma vez que esse é o nível de leitura frequentemente mais exigido para a realização de tarefas de um indexador.

3. Contatos para conhecer a disponibilidade do indexador

O próprio CIN contatou os indexadores para verificação de interesse pela pesquisa e disponibilidade para a entrevista, fazendo o agendamento da mesma.

4. Conversa informal com cada sujeito

No dia de cada entrevista foi realizada uma conversa informal com cada sujeito antes da realização Protocolo Verbal. Foram mencionados os objetivos da pesquisa mostrando sua importância para o desenvolvimento da área, evidenciando-se a preocupação de manter a identidade de cada um dos sujeitos oculta com o propósito da não comprometer os dados e deixá-los a vontade durante a realização da tarefa.

5. Familiarização com a tarefa do "ThinkAloud" (Pensar Alto) através de "Instruções aos Sujeitos"

Antes da aplicação do Protocolo Verbal ou Pensar Alto, foi feita uma familiarização com a tarefa utilizando um texto contendo "Instruções aos Sujeitos", elaborado com o propósito de descontrair e ao mesmo tempo apresentar procedimentos que poderão auxiliar o sujeito no desempenho da tarefa.

6. Gravação do "pensar alto" durante a leitura do texto-base

Antes de começar a gravação, foi entregue ao sujeito o texto-base lembrando que é preciso pensar alto durante toda leitura e exteriorizar seus processos mentais, procurando esquecer a presença da pesquisadora que estará presente apenas com o intuito de lembrar que é preciso pensar alto e também controlar o gravador.

7. Entrevista retrospectiva

Logo após a aplicação do protocolo verbal, foi feita uma entrevista retrospectiva com cada sujeito com o objetivo de esclarecer alguns pontos considerados obscuros pela pesquisadora.

8. Transcrição dos Protocolos Verbais dos sujeitos

As transcrições foram feitas de maneira a destacar a compreensão dos sujeitos, suas dúvidas, equívocos, identificação e seleção de termos.

Com a gravação do "Pensar Alto" durante a leitura e entrevista retrospectiva dos quatro sujeitos indexadores entrevistados, foi possível a transcrição literal completa em texto escrito e, em seguida, a análise dos dados coletados.

4 Resultados e discussão da observação das estratégias do indexador na leitura documentária

Considerando-se que as estratégias metacognitivas listadas por Brown (1987) e associadas à Norma são as mais representativas para a leitura com fins documentários, a discussão dos resultados obtidos realizou-se a partir dos seguintes aspectos:

I. Sequência de operações (aspectos) da leitura de sujeitos;
II. Manter em mente o seu objetivo de representar o texto para futura recuperação;
III. Exploração de seu conhecimento de estruturas textuais;
IV. Identificação de aspectos importantes da mensagem para seleção de conceitos;
V. Domínio da linguagem do sistema como conhecimento prévio.

Sequência de operações (aspectos) da leitura dos sujeitos

A sequência de operações, sintetizada a partir dos dados coletados na entrevista, dos quatro sujeitos, revelou pela análise documentária, que dos 10 tipos de operações consideradas indicativas de estratégias metacognitivas, o sujeito 1 utilizou 8 tipos, o sujeito 2 utilizou 6, o sujeito 3 utilizou 7 tipos e o sujeito 4 utilizou 4 tipos. Com relação ao padrão de sequências de operações, observou-se que, com exceção do sujeito 4 que repetiu somente 2 operações mais de uma vez, os outros sujeitos possuem um padrão de sequência, mesmo não sendo semelhantes ou cons-

tantes. Observa-se um ponto divergente entre os padrões: a operação 5 "Associação com linguagem" foi utilizada somente pelo sujeito 1. Essas constatações revelam que: a leitura feita pelos sujeitos é metacognitiva; os indexadores se utilizam muito mais do conhecimento prévio do assunto do que da linguagem documentária do sistema para a identificação de conceitos; há o domínio da estrutura textual de um artigo para ir direto às partes do texto em que foram identificaram aspectos importantes da mensagem e eles seguem a metodologia descrita no Manual INIS.

Manter em mente o seu objetivo de representar o texto para futura recuperação

De acordo com a análise das sequências da operações dos 4 sujeitos a operação "Explicitação dos objetivos da leitura" foi utilizada em 4 momentos pelo sujeito 1, em 3 momentos pelo sujeito 2, em 1 momento pelo sujeito 3 e em nenhum momento pelo sujeito 4. Como objetivo da leitura do indexador é representar o texto para futura recuperação pelo usuário do sistema, é extremamente necessário a manutenção deste objetivo em mente para identificação e seleção de conceitos que possam, realmente, expressar o conteúdo do documento, o que não ficou explícito durante a leitura do sujeito 4.

Exploração de seu conhecimento de estruturas textuais

Os sujeitos 1, 2, 3 foram os que melhor exploraram a estrutura textual, seguindo as recomendações do Manual INIS.O sujeito 1 explorou 6 partes da estrutura textual e dos 9 termos selecionados,6 são iguais ao INIS Thesaurus. O sujeito 2 explorou 4 partes da estrutura textual, selecionando 6 termos dos quais 4 são descritores. O sujeito 3 também explorou 4 partes da estrutura textual e selecionou 8 termos dos quais 6 são descritores. Jáo sujeito 4 explorou 3 partes da estrutura textual e selecionou 4 termos dos quais 3 são descritores.Os resultados demostram que os sujeitossão muito proficientes, explorando de forma eficaz a estrutura textual resultando a identificação de conceitos. Isto mostra que o domínio da estrutura textual facilita a exploração de modo a garantir a estratégia de identificação de conceitos.

Identificação de aspectos importantes da mensagem para seleção de conceitos

Na análise das partes da estrutura textual em que foram identificados conceitos, os resultados demonstram que o sujeito 1, além de ter sido o que mais explorou toda a estrutura textual, também utilizou o maior número de partes para identificação. Com relação ao sujeito 2, que também explorou toda a estrutura textual, observou-se que, para identificação de termos, limitou-se ao Título, Resumo, Introdução e Conclusão. O sujeito 3 explorou a estrutura textual, utilizando para identificação de conceitos, Introdução, Figura 1, Resumo e um subtítulo. O sujeito 4 limitou-se somente ao Resumo. A observação da sistemática de identificação de conceitos na leitura dos sujeitos demonstrou o uso de diferentes estratégias mas sem uma sistemática constante, ou seja, o sujeito 1 limitou-se a grifar conceitos que reconhece por associações com linguagem do sistema durante a leitura. Já o sujeito 2 relaciona partes importantes da estrutura textual com os termos identificados. O sujeito 3 somente grifa as palavras que reconhece como chave. O sujeito 4 não explicitou a utilização de nenhuma sistemática para identificação de conceitos. Os resultados dos sujeitos 1 e 2 combinam as estratégias de identificação de conceitos com o domínio da linguagem do sistema e da estrutura textual e conhecimento prévio do assunto, garantindo que os termos extraídos sejam mais representativos e ao mesmo tempo, compatíveis com a linguagem de recuperação.

Domínio da linguagem do sistema como conhecimento prévio

A verificação da compatibilidade dos termos selecionados pelos sujeitos com a linguagem do sistemaresultou em que , do total de 23 termos selecionados pelos sujeitos e pelo autor, 13 são descritores, 7 seriam adaptados, e 3 não foram encontrados. Correspondendo a 86,95% de compatibilidade com o sistema, conclui-se que os termos selecionados são **extremamente compatíveis.** Considerando-se que o indexador **é especialista**, o domínio do assunto se torna mais fácil e é utilizado como conhecimento prévio.

5 Considerações finais

Conforme os objetivos de identificar os procedimentos de leitura documentária utilizados para exame de documento e identificação de conceitos e examinar as estratégias de leitura usadas pelos indexadores para a identificação de conceitos que melhor representem o documento, foi realizado referencial teórico inicial sobre estratégias de leitura em Análise Documentária, bem como estudo de caso com o Centro de Informações Nucleares (CIN) da Comissão Nacional de Energia Nuclear (CNEN) para caracterização do indexador e levantamento das atividades de Análise Documentária.

A partir da síntese dos resultados obtidos, pode-se concluir que o indexador do Centro de Informações Nucleares é bastante estratégico e proficiente na realização da leitura para indexação.

Embora não sejam bibliotecários, os indexadores recebem treinamento específico do Sistema e contam com uma linguagem documentária bem estruturada e flexível e com Manuais de Referências para realização da indexação.

Sendo especialistas na sua área de atuação, os indexadores:

- Utilizam o domínio do assunto como conhecimento prévio e não a linguagem documentária do sistema, selecionando assim termos relacionados com sua área de atuação e extremamente compatíveis;
- Pouco utilizam a sequência de operações, pois quanto maior o conhecimento prévio do assunto menor a sequência de operações utilizada;
- Utilizam muito a exploração da estrutura textual;
- Não apresentam dificuldades durante a leitura;
- Seguem as recomendações do INIS para indexação.

A leiturado indexador do CIN é metacognitiva ao explorar toda a estrutura textual do documento, ao utilizar estratégias de leitura com os aspectos previstos por Brown (1980) e, principalmente, ao manter em mente

o objetivo da leitura – representar para recuperar. É vital para a indexação bem sucedida que o indexador realize estratégias de compreensão da leitura interagindo com o texto e tendo em vista o seu objetivo que é representar para que o usuário recupere.

Quanto à observação de estratégias de leitura utilizadas, ficou explícito que existem duas operações distintas utilizadas pelos indexadores durante a leitura: *Identificação de conceitos* e *Seleção de conceitos.* Isso esclarece que o indexador realiza as duas operações *durante* a leitura e não *após* a leitura. A Norma indica a realização das operações, mas não explicita o *momento* em que acontece, somente indicando que o estágio de Identificação de Conceitos deve ser feita após o exame do documento. De acordo com essa indicação, entende-se que a identificação de conceitos presentes no conteúdo do documento será realizada *após* a leitura e os conceitos serão, depois de identificados, representados por descritores de uma linguagem adotada pelo sistema de informação.

Tendo em vista os resultados obtidos na análise das estratégias através da técnica do Protocolo Verbal conclui-se que a presença de estratégias metacognitivas é resultado da compreensão da leitura e o uso maior ou menor de estratégias significa uma proficiência maior ou menor na compreensão do conteúdo do documento.

Referências

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 12676:* Métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, 1992.

BRASIL. Ministério da Ciência, Tecnologia e Inovação. Comissão Nacional de Energia Nuclear. Centro de Informações Nucleares. *Sobre o CIN*. 2015. Disponível em: <http://www.cnen.gov.br/centro-de-informacoes-nucleares/sobre-o-cin>. Acessoem: 11 jan. 2017.

BROWN, N. Metacognitive development and reading. In: SPIRO, R. J.; BRUCE, B. C.; BREWER, W. F. (Org).*Theoretical issues in reading comprehension.* New Jersey: L. Erlbaum, 1980. p. 453-481.

CAVALCANTI, C. R. *Indexação & tesauros:* metodologia & técnicas. Ed. Preliminar. Brasília: Associação dos Bibliotecários do Distrito Federal, 1978.

CAVALCANTI, M. C.; ZANOTTO, M. S. Introspection in Aplied Linguistics: meta-research on verbal protocols. In: BARBARA, L.; SCOTT, M. (Ed.). *Reflections on language learning.* Clevedon: Multilingual Matters , 1994. p. 148-156.

CHAUMIER, J. *Analisis y lenguajes documentales*. Barcelona: Mitre, 1986.

CHAUMIER, J. Indexação: conceito, etapas e instrumentos. Trad. José Augusto Chaves Guimarães. *Revista Brasileira de Biblioteconomia e Documentação*, São Paulo, v. 21, n.1/2, p. 63-79, jan./jun. 1988.

CINTRA, A. M. M. Estratégias de leitura em documentação. In: SMIT, J. W. (Coord.). *Análise documentária*: a análise da síntese. 2.ed. Brasília: IBICT, 1989. p. 29-37.

COHEN, A. D. Using verbal reports in research on language learning. In: FAERCH, C.; KASPER, G. (Ed.). *Introspection in second language research.* Clevedon: Multilingual Matters, 1987.p.82-95.

DIAS, E. W.; NAVES, M. M. L. *Análise de assunto*: teoria e prática. Brasília: Thesaurus, 2007.

ERICSON, K. A.; SIMON, H. A. Verbal reports on thinking. In: FAERCH, C.; KASPER, G. (Ed.). *Introspection in second language research.* Clevedon: Multilingual Matters, 1987.p.24-53.

FUJITA, M. S. L. *A leitura documentária do indexador:* aspectos cognitivos e lingüísticos influentes na formação do leitor profissional. 2003. 321f. Tese (Livre-Docência em Análise Documentária e Linguagens Documentárias Alfabéticas) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2003.

FUJITA**,** M. S. L. *A leitura em análise documentária.* 1998. 184 f. (Relatório de Pesquisa) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 1998.

FUJITA**,** M. S. L. *Leitura em análise documentária.* 2001. 185 f. (Relatóriode Pesquisa) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2001.

KATO, M. *O aprendizado de leitura.* São Paulo: Martins Fontes, 1987.

LANCASTER, F.W. *Indexação e resumos:* teoria e prática. 2.ed. Brasília: Briquet de Lemos, 2004.

NARDI, M. I. A. *As expressões metafóricas na compreensão de texto escrito em língua estrangeira.* 1993. 260 f. Dissertação (Mestrado em Lingüística Aplicada ao Ensino de Línguas) – Pontifícia Universidade Católica, São Paulo, 1993.

UNISIST. Princípios de indexação. *Revista da Escola de Biblioteconomia da UFMG,* Belo Horizonte,v.10, n.1, p. 83-94, mar. 1981.

VAN SLYPE, G. *Lenguages de indización:* concepción, contrucción y utilización en los sistemas documentales. Trad. Pedro Hípola e Félix de Moya. Madrid: Fundación Germán Sánchez Ruipérez, 1991.

Análise documental e representação da informação: aportes teóricos à utilização simultânea visando a recuperação da informação em Arquivos

Wendia Oliveira de Andrade
Dulce Amélia de Brito Neves

1 Introdução

No processo de organização da informação, temos o desenvolvimento de atividades e processos técnicos que irão culminar no acesso e uso por parte dos usuários, sejam eles internos ou externos à instituição na qual se encontra o arquivo, a biblioteca, o museu, ou qualquer unidade informacional que seja passível a busca de informações.

Ao longo desse processo informacional (organização, descrição, análise, representação e classificação), indagamos se existe um momento ímpar quando consideramos a efetiva recuperação da informação e, consequentemente, o atendimento à necessidade apresentada pelo usuário quando há uma lacuna (DERVIN et al., 2011) ou problema informacional (GROGAN, 2001).

Qual o momento que merece maior atenção no desenvolvimento das atividades ao lidar com o documento, seja ele impresso ou digital,

quando pensamos no momento de busca e recuperação documental e informacional por parte do usuário?

Da chegada do documento à unidade informacional e, por conseguinte,até seu usuário, temos um longo caminho de processos técnicos que, ao contrário do que se supõe, envolvem o desenvolvimento de tarefas minuciosas que exigem do profissional da informação[1] muito mais do que conhecimento operacional, pois o documento tem toda uma organicidade que não pode e nem deve ser desconsiderada ao longo desse caminhar dentro da unidade informacional. Aqui objetivamos refletir sobre o arquivo e seu documento, e as atividades que procuram analisar, representar e possibilitar sua recuperação.

O documento de arquivo é criado para atender determinado objetivo, seja ele administrativo, jurídico ou testemunhal. Além disso, esse documento tem em sua essência um caráter de prova, comprovação ou mesmo a representação do cumprimento de um ato, assim Bellotto (2008) nos apresenta a definição de documento de arquivo no contexto diplomático e a importância do Arquivo como unidade mantenedora de documentos tão importantes.

Essa perspectiva corrobora a importância da análise documental, relacionando-a com a diplomática arquivística, buscando elementos que comprovem a veracidade daquele documento ou, nas palavras da autora, a "comprovação de autenticidade" (BELLOTTO, 2008, p. 4). Sobre o formato, Camargo (2015, p. 16) destaca que "os documentos de arquivo são, por excelência, do gênero textual, ou seja, utilizam linguagem escrita para viabilizar certas ações e, a *posteriori*, para provar que essas mesmas ações se realizaram."

Schellenberg (2006, p. 41, grifo nosso) em sua definição clássica de documento de arquivo também ressalta o valor comprobatório quando afirma que:

> [...] no exercício de seus encargos legais ou em função das suas atividades e preservados ou depositados para preservação por aquela entidade ou por seus legítimos sucessores **como prova de suas funções, sua política, decisões, métodos, operações ou outras atividades, ou em virtude do valor informativo dos dados neles contidos.**

---

[1] Optamos por chamar de maneira mais genérica todos os profissionais inseridos nas unidades informacionais, em especial nos Arquivos, por compreender que, ainda no momento atual, temos a presença de bibliotecários, historiadores entre outros que não possuem a formação acadêmica em Arquivologia, não podendo chamá-los de arquivistas no sentido restrito da palavra, apesar de considerarmos muito sua experiência e contribuição para a área.

Nesse sentido, compreendemos que os documentos de arquivo possuem elementos que não podem ser desconsiderados quando analisados e representados, visando a sua recuperação futura, pois essas informações além de dar significação, apresentam valor informativo-comprobatório, mostrando que o documento de arquivo, além da forma, conteúdo ou suporte que se apresente, tem em seu teor informacional, elementos que, em determinado momento, fizeram-no ser essencial para corroborar um ato/ação no seu percurso de vida documental (valor primário).

Voltando nossa discussão especificamente ao contexto arquivístico na busca pela informação pelo usuário interno ou externo[2] (ANDRADE, 2014), temos vários aspectos que poderão incorrer em sucesso ou fracasso para a efetiva recuperação da informação arquivística. Dentre os fatores de sucesso, destacamos como imprescindíveis: a análise documental e a representação da informação.

Todas as informações contidas no documento são essenciais para que possamos compreender melhor o contexto documental e, dessa forma, melhor analisá-lo e representá-lo, tanto nos aspectos explorados pela análise diplomática e tipológica quanto na representação descritiva e temática do documento a ser estudado.

Em certos aspectos, são símiles a análise e a representação, pois tratam de descrever ao máximo o documento de arquivo, observando desde suas características físicas até o conteúdo informacional, passando pela espécie, tipo, fundo, gênero, entre outros elementos queconsideramos essenciais.

Devido à similaridade discricional em ambas as atividades, análise e representação, objetivamos aqui refletir, mesmo que teoricamente, sobre o que em essência apresentam essas teorias. Será que ambas podem ser utilizadas simultaneamente ao trabalhar com o documento de arquivo? Será que tratam da mesma coisa com nomenclatura distinta, ou apresentam pontos diferenciados que podem e/ou devem ser trabalhados em consonância?

---

[2] Podemos dizer que os usuários internos são aquelas pessoas que trabalham na mesma entidade mantenedora do arquivo. Eles conhecem a rotina administrativa, a documentação e todas as suas partes, sem necessariamente precisarem conhecer a rotina do Arquivo como unidade informacional, mas conhecem o documento e sabem que ele pertence à mesma instituição que ele. Por usuário externo, compreendemos aquele sujeito que não está vinculado à instituição, nem administrativa nem metodologicamente, ou seja, não existe nenhum tipo de contato com a unidade gestora que responda pela unidade informacional. Aqui entram os pesquisadores de maneira geral e os cidadãos comuns. (ANDRADE, 2014).

A partir de tais questionamentos, refletimos sobre alguns trabalhos que apresentam ideias fundamentais sobre diplomática e sobre análise documental com suas variantes diplomática e tipológica, especialmente o trabalho de Bellotto (2008) e de Duranti (2015). No tocante à indexação ou representação da informação (temática e descritiva), citamos Lancaster (2004), Maiomone, Silveira e Tálamo (2011), Fujita (2009; 2003) entre outros.

Compreendemos a existência de excelentes trabalhos a respeito da temática escolhida para dissertar, entretanto, escolhemos aqueles que, de acordo com nossa perspectiva, apresentam conceitos e definições mais claras e completas, o que nos permitiu uma melhor reflexão para o desenvolvimento de um estudo ao qual se pretende relacionar duas temáticas que têm pontos convergentes e divergentes (análise documental e a representação da informação) com objetivo comum: **a recuperação da informação para o usuário de arquivo,** seja ele interno ou externo.

2 Análises diplomática e tipológica

Analisar um documento, assim como representá-lo, não é tarefa somente objetiva ou descritiva, mas investigativa. Ao seguirmos os preceitos básicos para a recuperação da informação, estamos considerando o aspecto físico (o suporte) e também elementos subjetivosno que se refere ao seu conteúdo informacional (representação).

Ao adentramos no universo diplomático[3] e começamos a analisar a relação entre **matéria** e **conteúdo, suporte** e **informação,** assim como os **elementos externos** e **elementos internos** (BELLOTTO, 2008) **forma física** e **forma intelectual**[4] (DURANTI, 2015), notamos que não é algo simples de ser observado, pois a estrutura física nos apresenta elementos que necessitam de descrição e até mesmo de investigação,assim como oteor informacional do documento.

[3] Com base nas discussões de Belloto (2008), podemos entender que a Diplomática no contexto arquivístico é olhar para o aspecto formal do documento, considerando seus produtores e sua finalidade. Considera-se a importância desse documento e toda a conjuntura ao qual ele se insere, desde a unidade informacional, os produtores, o objetivo da criação do documento, a efetivação de sua função, sem deixarmos de observar toda a sua relevância como instrumento de prova.

[4] A diplomática define **forma** como o conjunto das regras de representação utilizadas para enviar uma mensagem, isto é, como as características de um documento que podem ser separadas da determinação dos assuntos, pessoas ou lugares específicos aos quais se referem. A forma documental é tanto física quanto intelectual. (DURANTI, 2015, p. 197).

A análise documentária ou documental pode ser entendida como um campo pertencente à Diplomática Arquivística[5] que trata de todos os aspectos constitutivos do documento de arquivo. Documento esse que apresenta características essenciais de prova, reforçando a ideia de que todo documento de arquivo é criado com fins comprobatórios, e após o cumprimento do ato para o qual foi criado -Bellotto (2008) refere-se ao ato jurídico – seguindo sua rotina administrativa, passando por sua temporalidade em cada idade respectiva (corrente e intermediária) e, por fim, chega à sua destinação final seja a eliminação ou a guarda permanente.

Dentro dessa perspectiva de análise documental, temos a análise diplomática e a análise tipológica que podemos dizer que:

> [...] são aplicações práticas dos estudos teóricos e metodológicos da diplomática e da tipologia documental, áreas das ciências sociais documentárias que se concentram, respectivamente, no estudo formal do documento diplomático, quando considerado individualmente, e no estudo das relações com o contexto orgânico de sua produção e atuação dos enunciados do seu conteúdo, quando considerados dentro dos conjuntos lógicos denominados séries arquivísticas. (BELLOTTO, 2008, p. ix).

Com base nessa definição, compreendemos que tanto as análises diplomática quanto a tipológica fazem parte de um mesmo cerne de análise ao documento de arquivo. E tal como foram apresentadas por Bellotto (2008), necessitam, por parte do profissional da informação, de um conhecimento prévio de alguns elementos que constituem a Arquivologia. Com intuito de facilitar os pontos mais importantes a serem observados no documento quando do uso de cada análise, apresentamos os itens descritos pela autora para aplicabilidade das técnicas:

---

[5] A partir do final do século XX, a Diplomática passa a ser dividida em Diplomática histórica e Diplomática Arquivística. Essa última, a partir da década de 80, começa a auxiliar os arquivistas na difícil tarefa de compreender o processo de criação dos documentos contemporâneos. Dessa forma, destaca-se a apropriação do método diplomático pela Arquivística, que passa a vê-lo como uma ferramenta *si ne qua non* para o tratamento da informação no século XXI. (TOGNOLI; GUIMARÃES, 2009, p. 25).

Quadro 1 – Sequência de identificação documental nas Análises Diplomática e Tipológica

	Identificação Diplomática	Identificação Tipológica
1	Autenticidade relativo à**espécie**, ao **conteúdo** e à finalidade	**Origem/proveniência**
2	**Datação**	Vinculação à competência e funções da entidade acumuladora
3	**Origem/Proveniência**	Relação entre **espécie** e tipo
4	Transmissão/tradição documental	**Conteúdo**
5	Fixação do Texto	**Datação**

Fonte: Elaborado com base em Bellotto (2008).

Através de uma simples compilação dos passos elencados por Bellotto (2008), percebemos que existem elementos que estão nos dois tipos de análise documental, tanto na diplomática quanto na tipológica, que é o caso dos itens anteriormente grifados no quadro: **espécie, conteúdo, datação e origem ou proveniência.**

Essa constatação nos remete a muitas coisas importantes, dentre elas a que é o foco do nosso estudo: justamente a indissociação dos dois tipos de análises quando mediante a avaliação documental. Todas as etapas elencadas por Bellotto (2008) são de extrema importância no processo de identificação do documento, objetivando com isso encontrar pontos que venham a convergir comsua proveniência, organicidade, autenticidade e fidedignidade, princípios fundamentais da arquivologia segundo a própria Camargo e Bellotto (1996).

Ainda nessa análise, percebemos quão fortes são os aspectos diplomáticos (ou mesmo formais) exigidos ao documento de arquivo, evidenciando que a Diplomática, mesmo agregando novos valores e técnicas, ainda preserva sua força no tocante à originalidade, no sentido de genuíno e não falsificado, do documento.

Mediante o exposto no quadro 1, encontramos indícios de que ambas as análises podem ser realizadas simultaneamente pelo profissional

da informação, desde que este tenha em mente o que objetiva cada uma delas. Se em ambas encontramos espécie, conteúdo, datação e origem e proveniência, podemos então sugerir que sejam essas as primeiras informações a serem buscadas no processo de análise, assim como a identificação da espécie documental aparecendo nos dois métodos,mas com enfoques distintos.

Na análise diplomática, espécie se relaciona diretamente com autenticidade e, na análise tipológica, a partir da Arquivologia ou da Arquivística como chama Bellotto (2008), relaciona-se com o tipo documental, ou seja, a espécie documental está presente em ambas as análises documentais, entretanto assumindo relações informacionais diferentes em cada uma. No primeiro caso, na diplomática, temos a apresentação de elementos que comprovemque o documento pode ser considerado legítimo, como forma, marcas, assinaturas, carimbos etc. (DURANTI, 2015) e, no segundo caso, temos a proveniência falando mais forte, exigindo um olhar para o produtor do documento e para que fins sejam criadosdentro da unidade informacional.

Diferentemente da representação, a análise documentalexige do profissional da informação um maior conhecimento arquivístico, já que se refere a termos que são específicos da Arquivologia como espécie, tipo, gênero, fundo, organicidade, proveniência, entre outros que assumem características específicas ao contexto informacional do Arquivo.

Aqui se faz necessário esclarecer nossa compreensão de espécie e tipo[6] na Arquivologia, pois são definições que comumente se confundem ainda mais quando consideramos a possibilidade de se trabalhar a análise e a representação concomitantemente.

Por **espécie** podemos entender a maneira como se apresenta um documento, ou seja, a forma textual que ele tende a ser redigido ou apresentado, respeitando sempre a unidade mantenedora que poderá fornecer um padrão de apresentação; deve conter sempre determinados elementos que o caracterizem como tal, por exemplo: atas, alvarás, atos, contratos, de-

[6] Não apenas no tocante à espécie e ao tipo, mas também quanto ao uso de outras terminologias específicas, deparamo-nos com esse problema descritivo na Arquivologia, como exemplifica Rodriguez (2015) sobre tipologia e tipo; além de dissertar a título expositivo sobre a multiplicidade do termo título em vários instrumentos arquivísticos normativos e sua complexidade de identificar a qual "título" se refere à norma: título formal, atribuído, oficial ou sistemático?

clarações etc. Por **tipo**, entendemos a espécie tornando-se específica, pois o documento apresenta certa estrutura, porém, em seu interior documental, há sua designação informacional que o torna único, como alvará de funcionamento, ou uma declaração de rendimentos, entre outros tipos específicos.

Aqui entra em foco a importância do conhecimento prévio da unidade informacional por parte do profissional que está no arquivo, para que, dessa form, haja uma melhor compreensãodas "relações orgânicas entre aquelas atividades entre si, entre elas e os documentos por ela produzidos/acumulados e entre os documentos em si" (BELLOTTO, 2008, p. 73).

Se pudéssemos, então,listar ou mesmo comparar os elementos a serem identificados em cada análise documental, seja ela diplomática ou tipológica, de acordo com o apresentado por Bellotto (2008), já estando de posse do documento e com arquivo, e tendo feita a primeira identificação documental, teríamos na análise diplomática três momentos:

1. **Protocolo Inicial:**

   Inovação;

   Titulação;

   Direção ou endereço;

   Saudação;

2. **Na parte Textual:**

   Preâmbulo;

   Notificação;

   Exposição;

   Dispositivo;

   Sanção;

   Corroboração ou cláusulas finais.

3. Protocolo final:

Subscrição/assinatura;

Datação;

Precação.

No tocante à análise tipológica, apresentam-se dois tipos que compilamos no quadro abaixo:

Quadro 2 – Análises Tipológicas a partir da Arquivologia

Análise Tipológica (Arquivistas Municipais de Madri)	Análise Tipológica (Modelo de Louise Gagnon-Arguin)
Tipo: espécie documental + atividade concernente	Tipo: denominação de tipo documental e sua categoria
Código: plano de classificação	Contexto de Criação
Entidade Produtora/Acumuladora	Definição
Atividade	Conteúdo
Destinatário	Condições de Validade
Legislação	Funções
Tramitação	Conservação
Documentos básicos	Autoria Responsável
Ordenação	Documentos Conexos
Conteúdo	Informações Complementares
Vigência	Leis
Prazos	

Fonte: Elaborado com base em Bellotto (2008).

Com base no exposto, identificamos que a análise tipológica aproxima-se mais do que comumente encontramos na literatura sobre a representação da informação, não obstante, ela é bastante específica à Arquivologia, pois busca identificar elementos de teor documental e sua relação com a entidade produtora e para que fins o documento foicriado. Assim como Bellotto (2008) ressalta, também compreendemos que os modelos de instrumentos de análise podem ser adaptados e adequados aos arquivos e aos seus documentos. Não há uma regra ou engessamento para realizar as análises, mas um roteiro para o desenvolvimento de tais procedimentos.

Sobre análise diplomática e tipológica, ambas podem ser consideradas complementares, já que uma encarrega-se de identificar elementos estruturais e até mesmo específicos, de forma e a outra se remete aos elementos informacionais do documento junto à entidade mantenedora e ao seu cumprimento de função, respectivamente;cada análise observará elementos que irão auxiliar não somente a identificação e autenticidade documental, mas também sua relação entre forma/conteúdo informacional, o que facilitará o processo representacional temático-descritivo.

3 REPRESENTAÇÃO DESCRITIVA E TEMÁTICA

Como vimos anteriormente, a análise documental, tanto diplomática quanto tipológica, trata de explorar o que Duranti (2015) chama de elementos extrínsecos e intrínsecos de forma de um documento de arquivo, buscando identificar elementos descritivos quanto ao suporte e ao conteúdo documental, elementos esses que requerem conhecimento prévio no tocante à Arquivologia, não explorando a informação de maneira mais ampla.

Partindo da premissa de que os elementos extrínsecos são "considerados aqueles que constituem o aspecto do documento e sua aparência externa" e os intrínsecos são "considerados como componentes integrais de sua articulação intelectual: o modo de apresentação do conteúdo do documento ou as partes que determinam **o teor do todo**" (DURANTI, 2015, p. 198-203, grifo nosso), constatamos que, apesar dessa última atividade nos remeter à ideia de "dentro de", "faz parte" ou "da natureza",os elementos tratam da forma do documento e não do seu teor informacional. Ou seja, as análises documentais buscam estudar o documento em profundidade arquivística, podendo outros aspectos informacionais, que seriam de grande ajuda na recuperação da informação, serem relegados a um segundo plano.

A leitura documental é feita e são identificados elementos que compõem literalmente o documento. É uma atividade que objetiva descrever o documento, isto é, dissecar item a item para melhor estudá-lo e identificar o máximo de partes possíveis.

Entendemos que as análises diplomática e tipológica são essenciais para um primeiro contato com o documento de arquivo,

explorar e identificar elementos que o compõem e, a partir delas, podermos desenvolver as demais atividades, dentre as quais temos a representação da informação, já que esta busca não somente os aspectos formais e descritivos dos documentos de arquivo.

A representação da informação é uma atividade desenvolvida com fins de recuperação da informação (independentemente do seu usuário ou do suporte documental). Tal representação objetiva identificar de que trata o documento; retirar os principais conceitos e transformá-los em uma linguagem artificial passível de recuperação manual ou informatizada. Lancaster (2004) traz essa ideia sob a nomenclatura de indexação, enfatizando que essa tarefa pode ser realizada em duas etapas conhecidas por análise conceitual e tradução. Fazendo uso do exposto por Lancaster (2004, p. 9-18), análise conceitual "implica decidir do que trata um documento – isto é qual é o seu assunto" e, no tocante à tradução "envolve a conversão da análise conceitual de um documento num determinado conjunto de termos de indexação."

A análise conceitual é a retirada dos descritores, dos conceitos mais representativos do documento. A tradução seria transformar a linguagem natural, ou seja, a linguagem do autor, em uma linguagem de indexação: concisa, precisa e que em poucos termos represente a informação de tal maneira que seja possível à recuperação.

Para Lancaster (2004), essas etapas são essenciais para uma indexação coerente e com o máximo de representatividade do teor documental. Chama atenção ainda para dispositivos de precisão e indicadores de função, todos os componentes que auxiliam no procedimento de atribuir, em maior ou menor grau, a precisão dos termos para a recuperação da informação.

Para Maiomone, Silveira e Tálamo (2011, p. 28, grifo nosso),a representação da informação:

> [...] é um processo que pode utilizar linguagem específica uma vez que possui características próprias e seu principal objetivo é **proporcionar a comunicação entre os documentos e os usuários da informação.**

Sendo um dos seus principais objetivos, disseminar a informação aos usuários que dela necessitem, a representação da informação realizada pelo profissional da informação, precisa estar atento quanto à responsabilidade em tal atividade. "A principal característica do processo de representação da informação é a substituição de uma entidade linguística longa e complexa – o texto do documento – por sua descrição abreviada." (NOVELLINO, 1996, p. 38).

Representar a informação contida em um documento requer do profissional da informação conhecimento da instituição que mantém o arquivo (estrutura e funcionamento), da essência documental (de que trata e para que fins foi criado o documento) e de a quem se destina o uso (KOBASHI, 1996). A indexação ou representação da informação não ocorre sem a análise documental, principalmente quando tratamos de Arquivos. São etapas de um processo representacional que são trabalhadas de forma tão uníssona que é bem comum pensarmos tratar-se da mesma atividade.

Maiomone, Silveira e Tálamo (2011, p. 28, grifo nosso) chamam atenção para a mudança de nomenclatura de atividades pertencentes à égide da análise conceitual, como a descrição bibliográfica e mesmo a catalogação para "representação descritiva" que "representa as características específicas do documento, denominada descrição bibliográfica, que permite a **individualização do documento.**" Tal definição nos remete à similaridade das atribuições informacionais da análise diplomática, com sua busca por pontos de acesso de um documento de arquivo que o torna de certa forma único ou individualizado, fazendo uso da expressão utilizada pelas autoras.

Já quando partimos para a representação temática, esta irá dar ênfase à "representação dos assuntos dos documentos a fim de aproximá-los, tornando mais fácil a recuperação de materiais relevantes no que dizem respeito a temas semelhantes" (MAIOMONE; SILVEIRA; TÁLAMO, 2011, p. 28). Quando pensamos na representação temática,é o momento no qual podemos fazer uso de instrumentos que auxiliem a recuperação da informação como unitermos e tesauros, que só podem ser construídos após uma descrição documental e uma análise prévia, identificando assim os elementos mais expressivos daquele documento e como estes podem relacionar-se entre si.

Tanto a representação descritiva quanto a representação temática apresentam elementos que se complementam, assim como a análise diplomática e tipológica. Entretanto, no caso da representação, é muito provável que não se logre êxito se o profissional da informação tentar começar seu trabalho pela temática ao invés da descritiva.

No caso das análises documentais, compreendemos que estas podem ser feitas simultaneamente ou a escolha de uma não irá interferir no desempenho da outra. No caso específico da representação, o mesmo não acontece, pois a representação descritiva torna-se a base para o passo seguinte, a representação temática, pois a representação descritiva torna-se exploratória para com o documento, possibilitando desenvolver a seguir um aprofundamento da informação documental já trabalhada (com descritores, classificação e organização já realizados).

4 Análises e representações: como fazer uso?

Após refletimos sobre essas duas atividades e também após a execução delas no contexto arquivístico, compreendemos que são atividades distintas com tronco comum: a recuperação da informação e o atendimento ao usuário da informação.

A análise tanto diplomática quanto tipológica são técnicas mais específicas do que a representação, pois elas falam diretamente aos arquivistas e supõem implicitamente um conhecimento prévio da Arquivologia, traz em seu cerne a história da diplomática e a importância da fidedignidade documental, além de sempre defender o documento de arquivo como um documento criado para provar, testemunhar ou corroborar algo. Tais efeitos chamam ainda mais a responsabilidade de se buscar elementos que confirmem a autenticidade documental.

Segundo Rodriguez (2015, p. 174), é preciso estar atentos ao realizar as análises, pois:

> A sequência de ações definida para a identificação diplomática e a identificação tipológica constitui elemento fundamental de análise e sua comparação permite evidenciar a diferença da abordagem própria de cada método. É possível perceber que diversos dos elementos constantes na identificação diplomática repetem-se na identificação tipológica, consistindo a principal diferença a ordem de procedimentos adotados para a análise documental.

O processo de análise está repleto de elementos que são específicos no tocante à Arquivologia como vimos em Bellotto (2008): espécie, tipo, entidade produtora ou acumuladora, a legislação que subsidiou a criação do documento, o estado de conservação entre outros aspectos. Já quando pensamos na representação da informação, temos recursos como a análise conceitual e a tradução, e também a elaboração de vocabulários controlados que auxiliariam no processo de busca por uma linguagem mais específica do documento.

Assim, como fazer uso das duas técnicas?

Como estamos dentro do uma perspectiva arquivística, já temos nosso ponto de partida, uma vez quediferente do que comumente encontramos na Biblioteconomia, cada documento de arquivo possui sua individualidade[7]. Por mais que apresente semelhanças de espécie e tipo (BELLOTTO, 2008), é preciso primeiramente conhecer a entidade mantenedora do Arquivo. Isso nos dará muitas informações e indícios do teor dos documentos produzidos e recebidos por ela.

Quando, mediante o arquivo e seu contato direto com a documentação, a leitura é indispensável, obviamente existem arquivos que não permitem uma leitura mais aprofundada da documentação. Por isso mesmo, dá-se a análise conceitual ou de assunto, ou seja, a retirada dos principais conceitos do documento **(representação descritiva)**e concomitantemente a identificação arquivística de elementos que identificam o documento, em seus elementos extrínsecos (DURANTI, 2015) realizando assim a **análise diplomática** do documento.

Nesse sentido, Fujita et al. (2009, p.122) complementa que

> O processo de representação é dependente da etapa de análise de assunto por meio da identificação e seleção de conceitos, com vistas à "tradução" desses conceitos identificados e selecionados por meio de termos constituintes de uma linguagem documentária. O processo de representação mediante linguagem documentária conduzirá o bibliotecário indexador à escolha dos termos correspondentes àespecificidadeeexaustividadequealinguagem possui e, consequentemente, à especificidade e exaustividade do sistema.

[7] Duranti (2015, p. 197) assevera que "[...] apesar das diferenças quanto à origem, procedência ou data, todos os documentos apresentam formas bastante semelhantes para possibilitar a concepção de uma forma documental típica, ideal, mais regular e completa, com o propósito de examinar todos os seus elementos."

Feita essa identificação primeira, unindo a representação descritiva e a análise diplomática, temos a base informacional para realizar as demais atividades como corrobora Bellotto (2008, p. 30, grifo nosso):

> O entendimento da análise diplomática tem, entre outras vantagens, facilitara elaboração de resumos de documentos, uma vez que o documentalista, conhecendo-a, há de realizar uma leitura documental mais segura. Ademais, esse procedimento é pré-requisito para o da análise tipológica, embora ela possa ser feita independentemente. Aliás, a análise tipológica pode tornar eficiente o fazer arquivístico em múltiplos aspectos, como a **avaliação, classificação e descrição.**

Dando prosseguimento ao trabalho de análise e representação para com o documento de arquivo, seguimos para a análise tipológica que, como apresentada por Bellotto (2008), pode ser adequada ao documento a ser trabalhado. A autora apresenta dois modelos de **análise tipológica** dos *Archiveros Municipales de Madrid* e da Gagnon-Arguin,os quais qualquer profissional da informação pode ter como parâmetro e criar seu próprio modelo a ser preenchido, adequando à realidade da sua unidade informacional e do seu fundo documental.

Ao longo de todo este trabalho, não deixamos de pensar na complexidade enfrentada pelo profissional da informação na busca por conceitos ou ideias que representem o teor documental, pois, sem dúvida, é um trabalho que exige dele não apenas objetividade, mas muita coerência ao desenvolvê-las, como Fujita (2003, p. 71) bem descreve:

> [...] no momento em que o indexador está lendo e procurando identificar e selecionar conceitos para a determinação do assunto do documento, está objetivando encontrar o assunto que lhe é familiar devido à sua prática de indexação e também definir o que pode interessar ao usuário do sistema de informação.

Lancaster (2004, p. 17) chama-nos a atenção para a melhor maneira de se representar a informação visando a sua recuperação. O autor faz então a seguinte explanação: "Se aceitarmos que a indexação é mais eficiente quando se orienta para a necessidade de determinados grupos de usuários, a função do indexador será prever os tipos de pedidos para os quais determinado documento será [...] útil."

Não obstante, Lancaster (2004) chama a atenção para o fato de que não é tão simplista o objetivo da indexação ou representação da informação, pois além de propiciar meios para que o usuário obtenha a informação que busca e saber de que trata o documento, tem a questão de como alimentar de forma correta o sistema de recuperação da informação que fará toda a diferença no sucesso ou fracasso na busca informacional. O autor ainda relata que o uso de uma linguagem muito específica para representar determinados documentos talvez não seja a melhor opção para a sua recuperação, por isso entendemos que as análises diplomática e tipológica sejam de extrema importância. Mas se foremas únicas utilizadas para descrever um documento visando a sua recuperação futura, provavelmente apenas os usuários que dominam a linguagem espeífica da Arquivologia irão efetivamente recuperar a informação descrita com base nos conceitos por elas escolhidos nas análises diplomática e tipológica.

Assim, durante a realização conjunta das análisesdocumentais e da representação descritiva, já podemos ter uma visão mais completa e contextualizadado documento, pois já houve uma investigação profunda de vários aspectos, tanto do documento quanto da unidade informacional a qual ele pertence. A partir desse ponto, já se pode pensar na construção de instrumentos elaborados com base na **representação temática**, ou seja, quando já ocorreu uma descrição documental e a escolha dos termos que mais representam o documento.

Esses instrumentos, como unitermos e tesauros, buscam trabalhar os conceitos escolhidos nas etapas anteriores, fazendo uso de uma linguagem específica, visando uma melhor recuperação da informação pelos usuários internos e externos ao arquivo, chegando, assim, ao grande objetivo da representação e da recuperação da informação: possibilitar o acesso e uso.

5 CONSIDERAÇÕES FINAIS

A descrição documental, sua análise e sua representação são elementos de grande importância quando pensamos na recuperação e uso da informação no contexto dos arquivos e de outras unidades informacionais.

Na literatura, encontramos normas que auxiliam o processo descritivo dos documentos de arquivo, mas sempre chamando a atenção para

o contexto no qual o documento está inserido. A norma internacional de descrição arquivística, a ISAD(G), por exemplo, apregoa que o objetivo da descrição arquivística:

> [...] é identificar e explicar o contexto e o conteúdo de documentos de arquivo a fim de promover o acesso aos mesmos. Isto é alcançado pela criação de **representações precisas e adequadas e pela organização dessas representações de acordo com modelos predeterminados**. Processos relacionados à descrição podem começar na ou antes da produção dos documentos e continuam durante sua vida. Esses processos permitem instituir controlesintelectuaisnecessáriosparatornarconfiáveis,autênticas, significativas e acessíveis descrições que serão mantidas ao longo do tempo. (CONSELHO INTERNACIONAL DE ARQUIVOS, 2001, p. 11, grifo nosso).

Com base no exposto na ISAD(G) (2000), especificamente no trecho ao qual demos ênfase, podemos entender as "representações precisas" como as análises diplomática e tipológica, sua organização e sua caracterização como usoda representação descritiva e temática. Apesar de as nomenclaturas de tais processos não aparecerem de forma explícita na referida Norma, entendemos o quão importantes são essas atividades analítico-representacionais.

Existem instrumentos como a própria ISAD(G) (2000) e mesmo a Norma Brasileira de descrição arquivística (NOBRADE, 2006) e suas 8 áreas com seus 28 elementos no tocante à descrição, que podem ser utilizadas como ferramentas no mister de desenvolvimento das atividades relativas à descrição documental não apenas no tocante ao suporte, mas também no que refere-se ao seu teor informacional.

Não nos faltam instrumentos para auxiliar o desenvolvimento dessas atividades tão importantes para a futura recuperação, mas o que chamamos a atenção especificamente nesse apontamento é que a análise documental e a representação, mesmo com objetivos similares, não se tratam da mesma coisa. Enquanto a análise acaba por tonar-se exaustiva quanto à forma, à espécie, ao tipo, à procedência, aos elementos gráficos entre outros, a representação em um momento seguinte busca tornar essas informações operacionalizáveis quanto à busca, faz com que todas as informações anteriormente exploradas possam ser selecionadas, organizadas

e até dispostas em um sistema de recuperação, facilitando seu acesso e uso. Não adentramos no aspecto linguístico e terminológico, apesar de entender que ele esteja diretamente ligado à questão descritiva e representativa. Isto porque a complexidade desse aspecto exigiria um estudo com ênfase apenas neste ponto. Como nosso objetivo é de dissertar e refletirsobre o desenvolvimento das atividades em conjuntocom a análise documental e com a representação para uma melhor recuperação da informação, não nos dedicamos especificamente à Terminologia.

Tanto as referidas análises quanto a representação da informação no contexto da Arquivologia podem ser realizadas conjuntamente. Sugerimos, para tanto, começar pela análise diplomática, seguida da análise tipológica que dá subsídios para a representação descritiva, e após a obtenção e organização desses elementos, passar à representação temática, ficando a critério do profissional da informação assim como da sua unidade gestora, a criação de instrumentos que auxiliem a recuperação da informação.

O que não pode deixar de ser observado no desenvolvimento dessas atividades é: a entidade que mantém o arquivo; sua dinâmica de funcionamento, para quem ela trabalha e para que fins e, só a partir dessas observações, tanto a análise quanto a representação irão compreender ao máximo o teor informacional dos documentos criados e também recebidos pela instituição. Assim, antes mesmo do desenvolvimento de qualquer atividade no arquivo, voltemo-nos sempre a Capurro (2003) quando nos indaga: informação - para quem?

Referências

ANDRADE, W. O. *Usuários da Informação Jurídica:* quem são e como funciona o fluxo informacional no Arquivo da Justiça Federal da Paraíba (JFPB). 138 f. 2014. Dissertação (Mestrado em Ciênciada Informação) — Universidade Federalda Paraíba, João Pessoa, 2014.

BELLOTTO, H. L. *Diplomática e tipologia documental em arquivos.* 2 ed. ver. eampl. Brasília, DF: Biquet de Lemos/Livros, 2008.

CAMARGO, A. M. A. Sobre as espécies e tipos documentais. In: INSTITUTO FERNANDO HENRIQUE CARDOSO. *Dar nome aos documentos:* da teoria à prática. São Paulo : Instituto Fernando Henrique Cardoso, 2015. p. 14-30.

______. ; BELLOTTO, H. L. *Dicionário de Terminologia Arquivística.* São Paulo: Secretaria da Cultura, 1996.

CAPURRO, R. *Epistemologia e Ciência da Informação.* In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5, 2003. Belo Horizonte, 2003. Anais eletrônicos... Belo horizonte: ENANCIB, 2003. Apresentação oral. Disponível em: <http://www.capurro.de/enancib_p.htm>. Acesso em: 20 nov. 2016.

CONSELHO NACIONAL DE ARQUIVOS (Brasil). *NOBRADE:* Norma brasileira de descrição arquivística. Rio de Janeiro: CONARQ, 2006. Disponível em: <http://www.conarq.arquivonacional.gov.br/Media/publicacoes/nobrade.pdf>. Acesso em: 20 jul. 2012.

CONSELHO INTERNACIONAL DE ARQUIVOS. *ISAD(G):* Norma geral internacional de descrição arquivística. 2. ed. Rio de Janeiro: Arquivo Nacional, 2001. 119 p. (Publicações técnicas, 49). Disponível em: <http://www.conarq.arquivonacional.gov.br/Media/publicacoes/isad_g_2001.pdf>. Acesso em: 20 nov. 2016.

DERVIN, B. et al. *Sense-Making as Methodology for Spirituality Theory, Praxis, Pedagogy, and Research.* In:FIRST GLOBAL CONFERENCE ON SPIRITUALITY IN THE 21ST CENTURY, 1, 2011. Praga, República Checa: Inter-Disciplinary.net, 2011. Disponível em:<https://www.inter-disciplinary.net/wp-content/uploads/2011/02/Dervin-paper1.pdf>. Acesso em: 24 nov. 2016.

DURANTI, L. Diplomática: novos usos para uma antiga ciência (Parte v). *Acervo- Revista do Arquivo Nacional,* Rio de Janeiro, v. 28, n. 1, p.196-215,jan./jun. 2015. Disponível em: <http://revista.arquivonacional.gov.br/index.php/revistaacervo/article/view/600>. Acesso em: 14 nov. 2016.

FUJITA, M. S. L. (Org.) et al. *A indexação de livros:* a percepção de catalogadores e usuários de bibliotecas universitárias. Um estudo de observação do contexto sociocognitivo com protocolos verbais [*online*]. São Paulo: Editora UNESP; Cultura Acadêmica, 2009.

______. A identificação de conceitos no processo de análise de assunto para indexação. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, SP, v. 1, n. 1, p. 60-90, dez. 2003. Disponível em: <http://www.sbu.unicamp.br/seer/ojs/index.php/rbci/article/view/287/167>. Acesso em: 20 nov. 2016.

GROGAN, D. J. *A prática do service de referência.* Tradução de Antonio Agenor Briquet de Lemos. Brasília, DF: Briquet Lemos/Livros, 2001.

KOBASHI, N. Y. Análise documentária e representação da informação. *Informare:* Cadernos do Programa de Pós-Graduação em Ciência da Informação, v. 2, n. 2, p. 5-27, 1996. Disponível em: <http://basessibi.c3sl.ufpr.br/brapci/v/a/3190>. Acesso em: 25 Nov. 2016.

LANCASTER, F. W. *Indexação e resumos:* teoria e prática. 2. ed. Brasília: Briquet de Lemos, 2004.

MAIMONE, G. D.; SILVEIRA, N. C.; TÁLAMO, M. F. G. M. Reflexões acerca das relações entre representação temática e descritiva. *Inf. & Soc.: Est.,* João Pessoa, v. 21, n. 1, p. 27-35, jan./abr. 2011.

NOVELLINO, M. S. F. Instrumentos e Metodologia de representação da informação. *Inf. Inf.,* Londrina, v. 1, n. 2, p. 37-45, jul./dez. 1996. Disponível em: <http://basessibi.c3sl.ufpr.br/brapci/_repositorio/2010/05/pdf_0e3cc20139_0010458.pdf>. Acesso em: 25 nov. 2016.

RODRIGUEZ, S. M. T. Atribuir nomes a tipos, séries e unidades documentais: dialogando com Mariano Garcia Ruipérez. In: INSTITUTO FERNANDO HENRIQUE CARDOSO. *Dar nome aos documentos:* da teoria à prática. São Paulo: Instituto Fernando Henrique Cardoso, 2015. p. 158-181.

SCHELLENBERG, T. R. *Arquivos modernos:* princípios e técnicas. 6. ed. Rio de Janeiro: Fundação Getúlio Vargas, 2006.

TOGNOLI, N. B.; GUIMARÃES, J. A. C. A diplomática contemporânea como base metodológica para a organização do conhecimento arquivístico: perspectivas de renovação a partir das idéias de Luciana Duranti. In: CONGRESSO ISKO ESPAÑA. 9., 2009, Valencia. *Anais eletrônicos...* Valencia:Universidad Politécnica de Valencia, 2009. p. 23-35. Disponível em: <http://www.iskoiberico.org/wp-content/uploads/2014/09/38-47_Bolfarini.pdf>. Acesso em: 14 nov. 2016.

DO MODELO CONCEITUAL FRSAD PARA O ASSUNTO PELA LEITURA: PERSPECTIVAS PARA A REPRESENTAÇÃO DOCUMENTÁRIA

Brisa Pozzi de Sousa

1 Introdução

Percebe-se que uma área de conhecimento necessita de sedimentação que estruture tanto sua história quanto sua fundamentação teórica e, tal estratégia, ocupa também a Biblioteconomia e Ciência da Informação (BCI), temos sobre tudo ao promover discussões responsáveis por abarcar os aspectos referentes à representação de documentos, tanto no seu aspecto formal quanto node conteúdo, ou temático. Portanto, essa representação compreende atividade essencial de novos cenários intelectuais, pois o objetivo é disponibilizar o documento para recuperação e assimilação por parte dos usuários.

Após a invenção da imprensa por Gutenberg e as ilimitadas possibilidades de edição eletrônica de documentos, tornou-se humanamente impossível conhecer, na proporção em que surge, o número de publicações de determinada área, sendo impraticável organizá-las individualmente. Assim, decorre a necessidade da organização por conteúdo, de forma padronizada, conforme realizado nos catálogos das bibliotecas.

A análise de assunto dos documentos, de forma manual, é subsidiada pela leitura documentária e demanda a institucionalização de proces-

sos como o de resumir, representar com descritores e elaborar linguagens documentárias. Em consequência, o contexto temático da leitura sustenta a elaboração dos processos e produtos documentários possibilitando a busca e recuperação da informação.

Mas, por um longo período na história, a representação temática de documentos esteve atrelada às operações empíricas do profissional e ao seu "[...] 'bom senso', pelas quais se atribuía, a um texto/documento qualquer, uma ou várias palavras-chave destinadas a facilitar a recuperação, no momento da pesquisa, sob um dado tema." (CUNHA, 1990, p. 59).

O cenário perdurou até meados do século XX, pois além do período do bom senso, imbuído de talento, experiência e do hábito profissional (CUNHA, 1990; KOBASHI, 1994), houve também receptividade em relação à determinação do conteúdo do documento e sua nomeação serem constituídos por um processo altamente intuitivo (GUIMARÃES, 2003). Obviamente que pensar uma área somente pelo seu fazer e crer que determinada atividade seria consequência de aptidões pessoais específica se representações unívocas não contribui com avanços significativos para sua teorização e sedimentação.

Remetendo ao cenário da Catalogação e a sua vinculação como processo global de confecção de catálogos[1], a *International Federation of Library Associations and Institutions*[2] (IFLA) inseriu a modelagem conceitual para a representação da informação, promovendo uma nova forma de fundamentar a área por meio dos *Functional Requirements for Bibliographic Records*[3] (FRBR). No ano de 1998, a IFLA publicou o relatório sobre o modelo FRBR, que além de contemplar o aspecto de quem faz uso das bases de dados bibliográficas, também abrangeu conceitualmente o "[...] universo bibliográfico e independente de um código de catalogação específico." (MELO; BRÄSCHER, 2014, p. 103). Portanto, também independente da linguagem documentária adotada pelo sistema.

---

[1] O catálogo é a base de dados que possibilita a busca e recuperação de registros informacionais e, em seus primórdios, por exemplo, eram utilizados como ferramenta de inventário.Assim "[...] a única exigência feita a um catálogo de biblioteca era que revelasse os itens componentes de determinada coleção." (SHERA; EGAN, 1969, p. 11). Todavia, da simples função de inventariar, atualmente os catálogos constituem importante instrumento de busca e recuperação da informação.

[2] Federação Internacional de Associações e Instituições Bibliotecárias.

[3] Requisitos Funcionais para Registros Bibliográficos.

A fim de estruturar a representação dos documentos, vários instrumentos como código de catalogação, tabela de classificação, tesauro, dentre outros, são utilizados para padronização dos registros, tanto em relação ao aspecto temático, quanto ao de forma, facilitando o processo comunicativo entre quem produz os catálogos ou bases de dados e aqueles que as utilizam.

O modelo FRBR tem a finalidade de operar em nível conceitual os aspectos referentes às entidades que representam os dados dos documentos, que são os objetos de interesse para os usuários, deliberando o processo comunicativo. Além disso, a família FRBR[4] totaliza três modelos: o próprio FRBR, o *Functional Requirements for Authority Data*[5] (FRAD) e o ***Functional Requirements for Subject Authority Data***[6] (FRSAD).

Entretanto, em 21 desetembro de 2016, a seção de Classificação e Indexação da IFLA passou a ser denominada Seção de Análise de Assunto e Acesso[7]. No início de 2016, outro movimento de mudança é constatado, pois em 28 de fevereiro, a Seção de Catalogação (*Cataloguing Section*) anunciou a integração dos três modelos (FRBR, FRAD e FRSAD) em um único, denominado *FRBR-Library Reference Model* (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2016, *on-line*).

Nesse debate conceitual, ressalta-se a indagação do modelo FRSAD comportar direcionamentos sobre a leitura documentária, pois seu foco centra no aspecto temático e, consequentemente, seus relacionamentos.

Sendo assim, pretende-se analisar o modelo FRSAD e buscar-se-á caracterizá-lo em seus aspectos mais gerais, com o objetivo de traçar a existência de uma possível visibilidade em torno da leitura documentária.

---

[4] Adotar-se-ánomeação "família FRBR" quando houver necessidade de algumapontamento concomitante aos três modelos – FRBR, FRSAD e FRSAD.

[5] Requisitos Funcionais para Dados de Autoridade.

[6] Requisitos Funcionais para Dados de Autoridade Assunto.

[7] "We propose "Subject Analysis and Access" as the new name of our section. It is short and a good compromise among the different proposals identified so far. In addition to the end-user focus ("Access") we also include the role of the professional. "Analysis" refers to the work and thought that goes into determining "aboutness" (or "determining what a document is about") and then creating subject access – be it assigning classification or headings, developing thesauri, or devising tools and other automatic ways to achieve subject access."(ŽUMER; ROE, 2016, p. 21).

2 Do modelo para assunto ao assunto pela leitura

Em consequência a necessidade de recuperação temática, a análise de assunto, que visa identificar o conteúdo, compreende uma das etapas de maior importância no processo de indexação (NAVES, 1996), pois através dessa análise o indexador terá que expressar com padrões o conteúdo do documento e disponibilizá-lo no catálogo, em forma de informação documentária.

Um ponto central consiste no fato da análise de assunto ser precedida pela leitura documentária e ambas são consideradas etapas que constituem o trabalho intelectual do(a) profissional. Como tema de pesquisa no Brasil, o princípio das discussões sobreleitura documentária destaca-se nas investigações de Cintra (1989), Kobashi (1994), Fujita, Nardi e Santos (1998) e Fujita (1999)[8], sendo a última responsável em trabalhara referida temática por um período de quinze anos.

A ação de representar tematicamente é a solução mais adequada para recuperar informações organizadas por assunto, como por exemplo, nos catálogos ou bases de dados. Na perspectiva de consolidar tais questões, houve a apropriação do modelo conceitual proveniente da Ciência da Computação e, a IFLA, através da Seção de Catalogação[9] e, da antiga Seção de Classificação e Indexação[10], propôs a investigação do modelo Entidade-Relacionamento (E-R), a fim de priorizar as necessidades dos usuários dos catálogos (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 1998).

O modelo E-R, desenvolvido por Peter Chen, embasa estruturalmente a família FRBRe é utilizado para estruturar e projetar bases de dados relacionais, em oposição às bases de dados hierárquicas. É conceitual, pois implica na modelagem de coisas, processos ou abstrações objetivando sintetizar e sistematizar sistemas, teorias ou fenômenos com

[8] Professora Mariângela Spotti Lopes Fujita iniciou as pesquisas sobre Leitura em Análise Documentária em 1993,sendo essa sua principal linha de investigação até o ano de 2009. Trabalhou com a temática em sua tese de livre-docência na Unesp câmpus de Marília, além de possuir várias publicaçõese inúmeras orientações de pesquisas de TCC, mestrado e doutorado sobre o tema. Portanto, a referida professora é considerada um marco na formação de pesquisadores que hoje atuam profissionalmente em universidades e em colaboração com outros pesquisadores sobre a referida temática.

[9] *Cataloguing Section* - http://www.ifla.org/cataloguing

[10] *Section on Classification and Indexing*, atual*Subject Analysis and Access* - http://www.ifla.org/classification-and-indexing

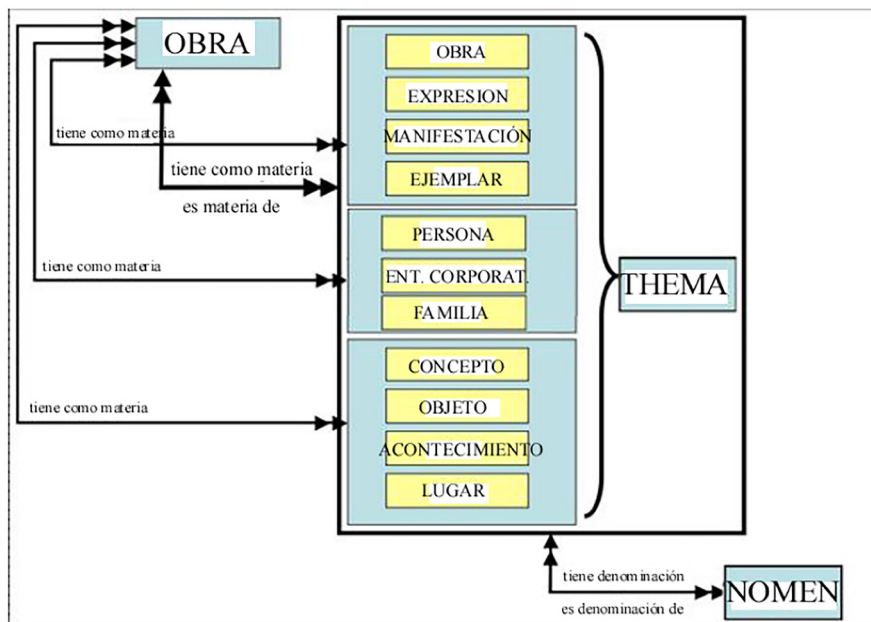
intuito de ser aplicável. Identifica-se no modelo E-R as entidades, os atributos e os relacionamentos. Considera-se entidade como algo, ou alguma coisa que pode ser distintamente identificada, relacionamento como a associação entre as entidades e, atributos, como características próprias de uma entidade (CHEN, 1990).

O modelo FRSAD representa uma análise das entidades que constituem os assuntos de uma obra, consistindo em determinar um marco que proporcione o entendimento dos registros bibliográficos e as respectivas autoridades de assunto com o objetivo de atender as buscas dos usuários (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2011).

Observando a família FRBR, composta pelos três modelos, totaliza-se o quantitativo de 11 entidades. O primeiro modelo, o FRBR, abrange as entidades dos produtos de trabalho intelectual ou artístico, sendo: obra, expressão, manifestação e item (constitui o grupo 1 de entidades). Já o segundo, o FRAD, compõe as entidades vinculadas aos responsáveis pela produção do conteúdo, disseminação e guarda das entidades do grupo 1, sendo: pessoa, entidade coletiva e família[11] (constitui o grupo 2 de entidades). E, o terceiro modelo, o FRSAD, determina as entidades que exercem os assuntos de uma obra, a saber: conceito, objeto, evento e lugar (constituem o grupo 3 de entidades) (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2011).

Todas as entidades do grupo 1 e 2 podem representar os assuntos de uma obra (grupo 3). Várias são as relações entre as entidades do registro bibliográfico e destacam-se as relações de assunto entre a entidade obra (grupo 1) com as demais entidades. Ilustra-se o fato com a figura que segue.

[11] Inserido a partir do modelo FRAD em 2009. Antes havia duas entidades nesse grupo: pessoa e entidade coletiva.

Figura 1 - Relações de assunto propostas pelas entidadesda família FRBR - grupo 1 (FRBR), grupo 2 (FRAD) e grupo 3 (FRSAD)

Fonte: International Federation of Library Associations and Institutions (2011)

Considera-se que os relacionamentos possuem a finalidade de apontar ligações entre as entidades, sendo tais ligações os recursos através dos quais os usuários poderão navegar, a fim de encontrar as conexões entre uma entidade e todas as outras que possam se relacionar a ela (TAYLOR, 2007).

É possível observar que outras duas entidades são destaque nos relacionamentos: *Thema* e *Nomen*. A primeira é definida como qualquer entidade utilizada como assunto de uma obra e, a segunda, como qualquer signo ou sequência de signos (caracteres alfanuméricos, símbolos, sons, etc) mediante os quais conhece, refere e aborda um *Thema* (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2011). O modelo ainda considera *Thema* uma superclasse de todas as entidades da família FRBR[12] e os termos matéria, tema e conceito são empregados como sinônimos e, nota-se que, equivocadamente, o modelo define todos como idênticos.

[12] "[...] *thema*es una superclase de todas las entidades FRBR [...]" (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2011, p. 13).

O modelo explica a escolha pelo latim para nomear as palavras "[...] *thema* (en plural *themata* o*Themas)* y *nomen* (en plural *nomina* o *nomens*), porque no tienenun significado pre-existente em nuestro contexto, son culturalmente neutros y no requieren traducción." (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2011, p. 13, itálico do modelo). Todavia, o latim pode sinalizar um resgaste do modelo não identificado por Aristóteles[13], que foi quem determinou as diretrizes da classificação do conhecimento por mais de dois mil anos. Além disso, segundo Marcondes (2007), na gênese de Aristóteles, a Filosofia Formal, ou seja, a Lógica aponta contribuição à Filosofia da Linguagem.

Nesse contexto, não é definido no modelo FRSAD fundamentação sobre o*Thema*. Ademais, várias questões permeiam o seu entendimento e, dentre elas, destacam-se:

- o modelo não define qual estrutura teórica é aporte da sua criação;
- não explora conceitualmente a questão do assunto, embora seja um modelo com enfoque para esse fim;
- reduz a estruturação do assunto entre*Thema* e *Nomen*;
- emprega os termos assunto, tema e conceito como sinônimos de *Thema,* sem nenhum escopo teórico para definição;
- vincula-se aos exemplos dos instrumentos elencadosno apêndice, como *La Faceted Application of SubjectTerminology* (FAST)[14] ao invés da discussão teórica das entidades que compõem o modelo;
- a produção científica que discuta o modelo FRSAD teoricamente e necessária e, sobretudo no Brasil, ainda é pequena; e

[13] "Ao valerem-se do latim como língua neutra, os criadores dos FRSAD pressupõem isenção quando, na verdade, não estão utilizando uma língua neutra. O latim é a língua utilizada por Aristóteles, por exemplo, e os preceitos filosóficos postulados por esse filósofo figuram como uma maneira de olhar a representação de assunto." (MILANI, 2014, p. 107).

[14] *La Faceted Application of Subject Terminology* (FAST) é uma adaptação da *Library of Congress Subject Headings* (LCSH).

- por último e não menos importante,indaga-se como o assunto podeser distintamente[15] identificado pelo viés da modelagem E-R.Aventar a possibilidade do conteúdo ser unívoco desfaz toda fundamentação teórica construída sobre a área temática e, portanto, nota-se ausência de discussão teórica sólida, sistematizada e abrangente sobre o modelo FRSAD.

Os pontos elencados não são novos e retomam problemas que, pelo modelo conceitual FRSAD, aparentemente, podem sinalizar como resolvidos, entretanto necessitam ser enfrentados teoricamente.

Além desses, notou-se que em nenhum momento a leitura do documento é indicada ou sinalizada como a maneira de atingir o assunto, pois a modelagem não substitui a análise. Ela estrutura os relacionamentos, entretanto o processo não é automático.

2.1 Perpassando o FRSAD

Em busca de delinear essa discussão, torna-se necessário recuperar alguns aportes teóricos e metodológicos que acenem para as necessidades de busca e acesso à informação dos usuários. "Estas necessidades variam de acordo com o domínio do saber, com o estado dos conhecimentos, com a natureza dos usuários e com seus objetivos." (GUINCHAT; MENOU, 1994, p. 28).

Nos últimos anos é possível observar o empenho no aprimoramento e na busca de avanços dos sistemas documentários, os quais são responsáveis em armazenar as informações documentárias, sendo estas

> [...] apreendidas, registradas e armazenadas em sistemas de informação documentária a fim de que sejam passíveis de recuperação e uso. As informações documentárias, portanto são unidades de representação, construídas sob uma forma e um conteúdo, a partir de decisões pautadas nos tipos de informação, nas áreas do conhecimento ou de atividade, na linguagem dos usuários e nos objetivos do serviço de informação, tornando explícito o propósito de um sistema de informação. (ORTEGA, 2008, p. 8).

[15] Importante salientar que conforme já citado (ver p. 3) com base em Chen (1990), a entidade é algo, ou alguma coisa que pode ser distintamente identificada. No dicionário consta como explicação para overbete "distinto": que não se confunde, perceptível, claro. (FERREIRA, 2010). Consequente, a questão do assunto de um documento ser inconfundível é simplesmente reducionista e retrógrado.

Dessa forma, as informações documentárias comportam operações no ciclo documentário, ou "cadeia documental" (GUINCHAT; MENOU, 1994) que se inicia com a produção de documentos, passa pela coleta, perpassa o tratamento ou organização daquilo que será armazenado para alcançar a recuperação, a disseminação e o uso da informação. "Estas operações são ligadas umas as outras, de tal forma que cada uma depende da que a precede, de acordo com a lógica do processo." (GUINCHAT; MENOU, 1994, p. 30).

A circulação de informações em um sistema documentário ou biblioteca, que comportem as etapas do ciclo documentário, ocasiona, por conseguinte, a

> [...] comunicação documentária e supõe a organização prévia da informação em categorias aptas a circular nas várias esferas da sociedade. Considera-se, assim, a presença de um sistema que elabora mensagens (o sistema de informação documentário) e o enunciatário (o usuário) que as recebe e as interpreta. A transferência de informações requer, portanto, a elaboração de mensagens (representações) que propiciem interpretações produtivas. (KOBASHI; TÁLAMO, 2003, p. 13).

É possível apreender que o conteúdo dos documentos se torna passível de socialização a partir da devida importância que é dada à sua organização. O foco não se resume na estocagem e centralização, mas no acesso às informações e, inclusive, nas formas que os documentos podem ser disponibilizados e acessados pelos usuários.

Passados aproximadamente 18 anos desde que o primeiro relatório final com o modelo FRBR foi estruturado, poucas pesquisas brasileiras são constatadas e, em relação ao modelo FRSAD, relativamente recente, pois teve versão publicada em 2010, cabe destacar duas pesquisas brasileiras que envolvem discussão sobre o tema, sendo a dissertação de Melo (2013) e a tese de Milani (2014).

O estudo de Melo (2013) investiga os tipos de relacionamentos entre conceitos que são tratados e representados nos sistemas de organização do conhecimento e indica a importância do uso da modelagem conceitual para a representação da informação. Aborda o propósito das tarefas dos usuários e a estrutura do modelo conceitual FRSAD, consti-

tuindo análise teórica do conceito (*thema*) e dos tipos de relacionamentos semânticos (*thema*-to-*thema*) estabelecidos no FRSAD, sob a perspectiva das abordagens positivista e pragmática do estudo do conceito. A autora conclui que a coexistência de abordagens teóricas em um modelo conceitual é uma relevante contribuição para a prática da representação temática da informação em sistemas de organização do conhecimento, argumento que encontra sua base de sustentação teórica nas pesquisas de Žumer, Zeng e Salaba (2012).

Já Milani (2014) caracteriza o fenômeno *bias*[16] na representação de assunto e aponta o poder que é instaurado pela construção de produtos por meio dos quais o usuário terá acesso à informação desejada. No entanto, de acordo com a autorasupracitada, há o risco de os substitutos documentais serem construídos com *biases* em dois contextos: negativo, quando deixam de incluir diversas características, conferindo pouca atenção aos grupos e temas, amparando-se em inclinações, desvios ou preconceitos; e positivo, onde prevalecem os direcionamentos ou vieses que visam garantir especificidade a determinadas comunidades usuárias.

O grupo de trabalho responsável pelo modelo FRSAD dispôs como objetivo estruturar conceitualmente as entidades do grupo 3 e como essas entidades se relacionam com os dados de autoridade assunto e as necessidades dos usuários. Além disso, o modelo intenciona auxiliar na avaliação do potencial de uso e compartilhamento internacional dos dados de autoridade assunto, tanto em bibliotecas como em outras instituições. O grupo de trabalho FRSAD, ou seja, *workinggroup* (FRSAD WG) aborda esse modelo no contexto dos demais, propostos pela IFLA, e desenvolvidos nos últimos anos (ZENG; ŽUMER, [2010?]).

Puente (2012) endossa que o modelo FRSAD pode ajudar no uso e desenvolvimento de dados de autoridade assunto e ressalta a independência deste com qualquer sistema ou contexto específico, permitindo assim o compartilhamento de dados e a interoperabilidade semântica.

[16] A autora destaca a opção por manter o termo original *bias(es)* em inglês, pois em língua portuguesa pode haver as conotações negativas de tendência, inclinação ou desvio e as conotações de direcionamento ou viés. Além disso, os instrumentos que embasam o processo de representação, como por exemplo, os tesauros, não são neutros e possuem valores morais inseridos. Com isso, os valores do profissional que realiza o processo de representação, assim como os valores do idealizador do instrumento refletirá nos produtos de representação.

Em uma busca complementar sobredocumentos que abordem especificamente o FRSAD, encontrou-se um livro publicado em inglês (ŽUMER; ZENG; SALABA, 2012).

As referidas três autoras apontadas são destaque no cenário internacional na abordagem do modelo FRSAD e, inclusive, são as editoras responsáveis pelo relatório final[17] do referido modelo, aprovado em junho de 2010, pelo Comitê da Seção de Classificação e Indexação da IFLA. Em decorrência, Žumer, Salaba e Zeng (2007) apontam que os estudos iniciais sobre o FRSAD, que então era denominado *Functional Requirements for Subject Authority Records* (FRSAR), foi constituído em abril de 2005 pelo grupo que estava envolvido com o desenho do modelo FRAD.

Ainda segundo as autoras supracitadas, todas as entidades dos três grupos têm potencial para ser o tema de uma obra, e o escopo do FRSAD consiste na abrangência de três propósitos: construir um modelo conceitual que relacione as entidades do grupo 3 no âmbito FRBR e do *aboutness*[18] das obras; fornecer definições claras e quadro de referência para relacionar os registros de autoridade assunto às necessidades dos usuários; e, por último, auxiliar na avaliação do potencial internacional de compartilhamento e uso dos dados de autoridade assunto, tanto dentro de bibliotecas como para além desse espaço (ŽUMER; SALABA; ZENG, 2007). Ressalta-se que os três propósitos também estão incluídos no relatório do modelo (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2010; 2011).

Considerando ainda Žumer, Salaba e Zeng (2007) é possível ratificar que o modelo FRSAD não possui sustentação teórica sobre a parte temática, objetivo principal que centra sua criação e função. Nota-se que prevaleceu um entendimento prático para a fundamentação, ou seja, do uso, quer de quem trabalhe ou pesquisa autoridade assunto em diferentes espaços, conforme é possível constatar:

---

[17] INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. *Functional Requirements for Subject Authority Data (FRSAD):* a conceptual model. 2010. Disponível em: <http://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>. Acesso em: 1nov. 2016.

[18] Um dos objetivos que o modelo se propõe é relacionar o *aboutness* das obras, fato que incita vários debates e direcionamentos teóricos, conforme estabelecido por autores que perpassam várias áreas, dentre elas a CI, a filosofia, a semiótica.Destaque para: Hutchins (1977); Hjørland (1992); Putnam (1958); Alves, Moraes e Almeida (2014); Guedes (2009); Moraes (2012).

> The FRSAR Working Group felt strongly that, in order to define user tasks, an actual user study was necessary, and two studies were therefore conducted. The first was a pilot study at the 2006 Semantic Technologies Conference (San Jose, California, USA). Most study participants were either creators of semantic tools, including controlled vocabularies, taxonomies and ontologies, or developers and managers of semantic technology systems. The second study was an international survey sent to information professionals throughout the world during the months of May-September 2007. Participants included authority record creators, vocabulary creators and managers, catalogers, metadata librarians, and reference librarians among others. Participants were asked to describe their work and their use of subject authority data in different contexts, including cataloging/metadata creation, subject authority work, and searching or helping others search bibliographic information. The results of these studies enriched our understanding of subject authority data use and informed and further confirmed the FRSAR user tasks[19]. (ŽUMER; SALABA; ZENG, 2007, p. 490).

Em continuidade há indicação de trabalhos apresentados em eventos pelas editoras do modelo FRSAD eaponta-se: Zeng e Žumer (2009a) que aborda mapeamento do modelo FRSAD e outros modelos abstratos, apresentado na ISKO UK *Conference*, em Londres; Zeng e Žumer (2009b) que indica introdução ao FRSAD e mapeamento com uso de SKOS e Žumer e Zeng (2015) publicação que indica aplicação do FRBR e FRSAD em sistemas de classificação. Observa-se na abordagem das autoras a vinculação do modelo FRSAD a instrumentos e não a discussões sobre possíveis aportes teóricos para fundamentação. Os trabalhos de 2009b e 2015 estão vinculados aos *Simple Knowledge Organization System* (SKOS) ou Sistemas de Organização do Conhecimento e, inclusive, o segundo a *Universal Decimal Classification* (Classificação Decimal Universal).

[19] O Grupo de Trabalho FRSAR, sentiu necessidade de definir as tarefas do usuário, um estudo real foi necessário e dois foram conduzidos. O primeiro foi um estudo piloto em 2006 durante a *Semantic Technologies Conference* (San Jose, Califórnia, EUA). A maioria dos participantes do estudo eram ou criadores de ferramentas semânticas, incluindo vocabulários controlados, taxonomias e ontologias, ou desenvolvedores e gestores de sistemas de tecnologia semântica. O segundo estudo foi uma pesquisa internacional enviada para profissionais da informação em todo o mundo durante os meses de maio a setembro de 2007. Entre os participantes foram incluídos criadores de registro de autoridade, criadores e gestores de vocabulário, catalogadores, bibliotecários de metadados e de referência, entre outros. Os participantes foram solicitados a descrever seu trabalho e a utilização de dados de autoridade assunto em diferentes contextos, incluindo catalogação/criação de metadados, o trabalho de autoridade assunto, e pesquisas ou auxílios a outras pessoas por buscas de informações bibliográficas. Os resultados destes estudos enriqueceu nossa compreensão do uso de dados autoridade assunto e informou e confirmou as tarefas do usuário no [modelo] FRSAR.(ŽUMER; SALABA; ZENG, 2007, p. 490, traduçãonossa).

3 A LEITURA DOCUMENTÁRIA

Nota-se que a indexação é utilizada para indicar o procedimento de catalogação de assunto em bibliotecas. O fato é decorrente do avanço e uso das tecnologias, que ocasionaram a transformação dos catálogos manuais em *on-line,* sendo estes considerados equivalentes às bases de dados (FUJITA, 2009).

Anterior ao processo de transferência e recuperação da informação tem-se que considerar o tratamento, que decorre da necessidade de representar os documentos, para disponibilizá-los à comunidade usuária de maneira tratada. Salvo contrário, o acervo de uma biblioteca estaria abnegado ao acaso, formando um conglomerado de informações dispersas.

Considerada a etapa que constitui o trabalho intelectual do indexador, o foco nessa seção percorre a abordagem da fase inicial do processo de indexação denominada análise de assunto. No entanto, também pode ser concebida por outras denominações e "o processo de extrair conceitos que traduzam a essência de um documento é conhecido como análise de assunto para alguns, análise temática para outros e ainda como análise documentária ou análise de conteúdo." (NAVES, 1996, p. 215).

Os estudos de Cintra (1989) fundamentaram a importância da compreensão do ato de ler para o cumprimento da atividade de indexação. Kobashi (1994) condiciona a importância da realização da leitura no contexto de análise documentária. Cunha (1990, p. 141, destaque da autora) explica que a análise do documento "[...] passa em primeiro lugar pela LEITURA do texto e sua análise [...]" sendo esta etapa o "momento deflagrador" da análise.

Com isso, observa-se um articulado movimento de interação entre leitor e texto a ser incorporado no sistema documentário, tendo a leitura do profissional o papel de apropriação do conhecimento, que segundo Moura (2004), remete a dois tipos específicos de apropriação textual:

> De um lado, temos a apropriação terminológica caracterizada pelo contínuo armazenamento, em enciclopédia particular, de uma lista de descritores aplicáveis às áreas de atuação do profissional. No decorrer do tempo, devido à familiaridade com a terminologia, tal leitor é capaz de incorporar novas informações aos sistemas, formular expressões

> de busca, estabelecer diálogos interdisciplinares sem, contudo, conseguir extrair as implicações teórico metodológicas desta ação. É uma habilidade que o bibliotecário desenvolve ao longo de suas atividades profissionais.
>
> Por outro lado, a apropriação conceitual ocorre de modo mais consciente. Neste caso, o profissional busca, no desempenho de suas atividades,conhecer de forma mais específica os conceitos e as interfaces que os descritores representam para as distintas áreas de conhecimento. Devido a esse fato, a representação de um dado item informacional realizada pelo bibliotecário ganha significação e a mediação passa a ter mais chance de efetivação. (MOURA, 2004, p. 165).

Na segunda forma de apropriação descrita, predomina a maneira mais consciente de assimilar as informações documentárias que são produzidas, sendo observável a ligação entre o profissional, o sistema de recuperação da informação e o usuário. A autora explica a importância de o bibliotecário atuar como um leitor-mediador, a fim de "[...] exercer a capacidade interpretativa sã e, às vezes, a paranóica, para melhor compreender a dinâmica do fenômeno informacional, o que exige dos profissionais uma postura multidisciplinar na realização do seu trabalho [...]." (MOURA, 2004, p. 167).

A fim de oferecer subsídios aos profissionais em relação à atividade de determinar o assunto do documento, foi criado os "Princípios de Indexação" pelo sistema internacional vinculado à Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO)[20], conhecido pela sigla UNISIST[21] (WORLD INFORMATION SYSTEM FOR SCIENCE AND TECHNOLOGY, 1981; FUJITA, 2003).

Em 1985, a publicação dos Princípios de Indexação ocasionou a criação da primeira norma internacional, a ISO 5963 e, no Brasil, somente 7 anos depois, em 1992, a Associação Brasileira de Normas Técnicas (ABNT) publicou a tradução, nomeando-a de NBR 12676/1992 (FUJITA, 2003).

Fujita (2003) investiga por meio da revisão de literatura, a identificação de conceitos a partir da leitura documentária, da tematicidade e das

---

[20] *United Nations Educational, Scientific and Cultural Organization.*

[21] *World Information System for Science and Technology.*

concepções de análise do assunto do documento, abordando, indicando e tendo como aparato além de outros autores e a NBR 12676/1992[22], os Princípios UNISIST.

O processo de indexação é definido por Princípios UNISIST (WORLD INFORMATION SYSTEM FOR SCIENCE AND TECHNOLOGY, 1981, p. 84), como sendo:

> [...] a ação de descrever e identificar um documento de acordo com seu assunto. [...] Durante a indexação, os conceitos são extraídos do documento através de um processo de análise, e então traduzidos para os termos de instrumentos de indexação (tais como tesauros, listas de cabeçalhos de assunto, esquemas de classificação, etc).

Também de acordo com os Princípios de Indexação (WORLD INFORMATION SYSTEM FOR SCIENCE AND TECHNOLOGY, 1981, p. 85), a operacionalização do processo decorre em dois estágios sendo: "- Estabelecimento dos conceitos tratados num documento, isto é, o assunto; - Tradução dos conceitos nos termos da linguagem de indexação."

Focando no texto referente aos Princípios, durante o estágio de determinação de assunto (1º estágio) é possível observar que a leitura documentária encontra-se subentendida em "partes importantes do texto [que] necessitam ser consideradas cuidadosamente" (WORLD INFORMATION SYSTEM FOR SCIENCE AND TECHNOLOGY, 1981, p. 86), pois mesmo para compreender as partes, é necessário realizar a leitura documentária.

De acordo com Fujita (2003, p. 64, destaque da autora) é possível constatar que a leitura está subtendida nos Princípios de Indexação, na frase grifada:

> Na *identificação de conceitos* (segundo estágio do estabelecimento de conceitos), o indexador, após o exame do texto, passa a abordá-lo de uma forma mais lógica a fim de selecionar os conceitos que melhor representem seu conteúdo. Para isso, recomenda que a identificação de conceitos seja feita obedecendo a um esquema de categorias existente na área coberta pelo documento, como por ex.: o fenômeno, o processo, as propriedades, as operações, o material, o equipamento, etc.

---

[22] Norma Brasileira responsável em subsidiar o processo de indexação. Entre os objetivos da referida Norma, encontra-se a fixação de condições exigíveis para a prática normalizada do exame de documentos, da determinação de seus assuntos e da seleção de termos de indexação (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992).

São nítidas as diferenças entre os Princípios UNISIST e o próprio modelo conceitual FRSAD, entretanto, ambos possuem a dimensão de terem sido articulados por entidades internacionais e, assim, carregam visibilidade ampliada. Nessa perspectiva, tais iniciativas são de grande importância para sustentar discussões sobre a finalidade e importância do processo de representação.

O fato de resgatar o UNISIST não se vincula a tecer comparações entre este e o modelo conceitual FRSAD, pois o primeiro possui uma proposta de princípios e, o segundo, da modelagem conceitual, entretanto ambos perseguemo mesmo resultado: a representação. Mas, nem de forma sublimar o modelo da IFLA aponta o subsídio da leitura para a estruturação dos assuntos.

4 Considerações finais

Através do exposto é possível perceber que a leitura documentária possui condições de permear o contexto dos modelos propostos pela IFLA, sobretudo pelo fato desses estarem em processo de discussão. Para tanto,se considera que uma nova dimensão precisa ser assumida pelo modelo, na qual o assunto do documento passe a ser desencadeado de forma relacionada à leitura documentária.

O reforço teórico para estabelecera representação temática torna-se uma possibilidade de derrubar o estigma do tratamento da informação documentária ser visto como resultado somenteda prática profissional, portanto sem fundamentação.

Independente das propostas que sustentem a discussão, o foco sempre será a disponibilização da informação documentária em catálogos ou bases de dados, a fim de facilitar a busca e a recuperação. Nesse entendimento, somente a partir da integração entre os escopos teóricos e práticos será possível consolidar a parceria entre leitura documentária, análise de assunto, representação e recuperação, aspectos tão necessários e caros a Biblioteconomia e Ciência da Informação.

Referências

ALVES, R. C. V.; MORAES, J. B. E.; ALMEIDA, C. C. Semiótica do discurso científico: um estudo sobre novas perspectivas para a análise documental de conteúdo. *Scire*, v. 20, n. 2. p. 55-59. jul./dic. 2014.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 12676:* métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, 1992.

CHEN, P. P. *Modelagem de dados:* a abordagem entidade-relacionamento para projeto lógico. São Paulo: MakronBooks,1990.

CINTRA, A. M. M. Estratégias de leitura em documentação. In: SMIT, J. W. (Coord.). *Análise documentária:* a análise da síntese. Brasília: IBICT, 1989. p.27-35.

CUNHA, I. M. R. F. *Do mito à análise documentária*. São Paulo: EDUSP, 1990. (Teses; v. 11).

FERREIRA, A. B. H. *Dicionário da língua portuguesa*. 5. ed. Curitiba: Positivo, 2010.

FUJITA, M. N. S. L.; NARDI, M. I. A.; SANTOS, S. A. S. A leitura em análise documentária. *Transinformação*, v. 10, n. 3, p. 13-31, 1998.

FUJITA, M. N. S. L. A leitura do indexador: estudo de observação. *Perspectivas em Ciência da Informação*, v. 4, n. 1, p. 101-116, 1999.

FUJITA, M. S. L. A identificação de conceitos no processo de análise de assunto para indexação. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v. 1, n. 1, p. 60-90, jul./dez. 2003.

FUJITA, M. S. L. O contexto da indexação para a catalogação de livros: uma introdução. In: FUJITA, M. S. L. (Org.). *A indexação de livros:* a percepção de catalogadores e usuários de bibliotecas universitárias. São Paulo: Cultura Acadêmica, 2009. p. 11-17.

GUEDES, E.G. F. *O conceito aboutness na Organização e Representação do Conhecimento*. 90 f. 2009. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, 2009.

GUIMARÃES, J. A. C. A Análise documentária no âmbito do tratamento da informação: elementos históricos e conceituais. In: RODRIGUES, G. M.; LOPES, I. L. (Org.). *Organização e representação do conhecimento na perspectiva da Ciência da Informação*. Brasília: Thesaurus, 2003. Série Estudos Avançados em Ciência da Informação. p. 100-117.

GUINCHAT, C.; MENOU, M. *Introdução geral às ciências e técnicas da informação e documentação.* Brasília: IBICT, 1994.

HUTCHINS, W. J. On the problem of "aboutness" in document analysis. *Journal of Informatics*, v. 1, n. 1, p. 17-35, 1977.

HJØRLAND, B. The concept of "subject" in information science. *Journal of Documentation*, v. 48, n. 2, p. 172-200, 1992.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. *Requisitos Funcionales de los Registros Bibliográficos:* informe final. [Madrid]: Ministerio de Cultura, 1998. Traducción Xavier Agenjo y MaríaLuisa Martínez-Conde. Disponível em: <http://www.ifla.org/files/assets/cataloguing/frbr/frbr-es.pdf>. Acessoem: 4 ago. 2014.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. *Functional Requirements for Subject Authority Data (FRSAD):* a conceptual model. 2010.Editors: Marcia Lei Zeng; MajaŽumer; Athena Salaba. Disponível em: <http://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>. Acessoem: 1 mar. 2016.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. *Requisitos Funcionales para Datos de Autoridad de Materia (FRSAD):* un modelo conceptual. España. 2011. Traducción: Pascual Jiménez Huerta, Lourdes Alonso Viana. Disponível em: <http://www.ifla.org/files/assets/cataloguing/frbr/frbr-es.pdf>. Acessoem: 4 ago. 2014.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS.Section Cataloguing. FRBR Review Group. *World-wide review of the FRBR-Library Reference Model, a consolidation of the FRBR, FRAD and FRSAD conceptual models.* 28 february 2016. Disponível em: <http://www.ifla.org/node/10280?og=54>. Acesso em: 29 set. 2016.

KOBASHI, N. Y. *A elaboração de informações documentárias*: em busca de uma metodologia. 1994. 195f. Tese (Doutorado em Ciências da Comunicação)– Programa de Pós-Graduação em Ciências da Comunicação, Universidade de São Paulo, São Paulo, 1994.

KOBASHI, N. Y.; TÁLAMO, M. F. G. M. Informação: fenômeno e objeto de estudo da sociedade contemporânea. *Transinformação*, Campinas, v. 15, n. 3, p. 7-21, set./dez. 2003. (Edição Especial).

MARCONDES, D. *Iniciação à história da filosofia*: dos pré-socráticos a Wittgenstein. Rio de Janeiro: Zahar, 2007.

MELO, M. A. F. *Contribuições das abordagens positivista e pragmática do estudo do conceito para o modelo conceitual FRSAD*. 2013. 133 f. Dissertação (Mestrado em Ciência da Informação)–Programa de Pós-Graduação em Ciência da Informação, Universidade de Brasília, Brasília, 2013.

MELO, M. A. F.; BRÄSCHER, M. Requisitos Funcionais para Dados de Autoridade Assunto (FRSAD): entidades, atributos e relacionamentos. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v. 12, n. 2, p.102-119, maio/ago. 2014.

MILANI, S. O. *Bias na representação de assunto*: uma discussão de oposições Bináriasnos Functional Requirements for Subject Authority Data (FRSAD). 2014. 134 f. Tese (Doutorado em Ciência da Informação)–Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2014.

MORAES, J. B. E. de. Perspectivas metodológicas para a identificação do aboutness em textos narrativos de ficção. *Scire*, v. 18, n. 2. p. 57-66. jul./dic. 2012.

MOURA, M. A. Leitor-bibliotecário: interpretação, memória e as contradições da subjetividade. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 9 n. 2, p. 158-169, jul./dez. 2004.

NAVES, M. M. L. Análise de assunto: concepções. *Revista de Biblioteconomia de Brasília*, Brasília, v. 20, n. 2, p. 215-226, jul./dez. 1996.

ORTEGA, C. D. Fundamentos da organização da informação frente à produção de documentos. *Transinformação*, Campinas, v. 20, n. 1, p. 7-15, jan./abr. 2008.

PUENTE, M. *Nuevos paradigmas em elcontrol de autoridades temáticas*: el modelo FRSAD y los sistemas de organizacióndelconocimiento. Buenos Aires: Consultora de Ciencias de laInformación, 2012. (Serie Documentos de Trabajo; Área Procesos Técnicos). Disponível em: <http://eprints.rclis.org/17399/1/DT_036.pdf>. Acessoem: 2 set. 2016.

PUTNAM, H. Formalization of the concept "about." *Philosophy of Science*, v. 25, n. 2, p. 125-130, 1958.

SHERA, J. H.; EGAN, M. E. *Catálogo sistemático*: princípios básicos e utilização. Brasília: EditôraUniversidade de Brasília, 1969.

TAYLOR, A. G. An introduction to functional requirements for bibliographic records (FRBR). In: TAYLOR, A. G. *Understanding FRBR*: what it is and how it will affect our retrieval tools. Westport: Green wood Publishing Group, 2007. p. 1-19.

WORLD INFORMATION SYSTEM FOR SCIENCE AND TECHNOLOGY. Princípios de indexação. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, v. 10, n. 1, p. 83-94, mar. 1981.

ZENG, M. L.; ŽUMER, M. Mapping FRSAD model and other abstract models. In.: ISKO UK Conference, 2009a, London. *Proceedings…* London: ISKO UK, 2009a. Disponível em: <http://www.iskouk.org/content/mapping-frsad-model-and-other-abstract-models>. Acessoem: 3 mar. 2016.

ZENG, M. L.; ŽUMER, M. Introducing FRSAD and Mapping it with SKOS and other models. In.: WORLD LIBRARY AND INFORMATION CONGRESS, IFLA GENERAL CONFERENCE AND COUNCIL, 75., 2009b, Milan. *Proceedings…* Milan: IFLA, 2009b. Disponível em: <http://www.ifla.org/past-wlic/2009/>. Acessoem: 3 mar. 2016.

ZENG; M. L.; ŽUMER, M. *Introducing FRSAD and mapping it with SKOS and other models.* [2010?]. Based on the work of the FRSAR Working Group. Disponívelem: <http://nkos.slis.kent.edu/FRSAR/introducingFRSAD.pdf>. Acessoem: 2 set. 2016.

ŽUMER, M.; SALABA, A.; ZENG, M. L. Functional Requirements for Subject Authority Records (FRSAR): a conceptual model of aboutness. In.: INTERNATIONAL CONFERENCE ON ASIAN DIGITAL LIBRARIES, 10., 2007, Hanoi, Vietman. *Proceedings…* Germany: Springer-Verlag Berlin Heidelberg, 2007. p. 487 - 492.

ŽUMER, M.; ZENG, M. L.; SALABA, A. *FRSAD:* conceptual modeling of aboutness. California: Libraries Unlimited, 2012.

ŽUMER, M.; ZENG, M. L. Application of FRBR and FRSAD to classification systems.In:INTERNATIONAL UDC SEMINAR, 2015, Lisbon. *Proceedings….* Würzburg: Ergon-Verlag, 2015. Edited by Aida Slavic e Maria Inês Cordeiro.

ŽUMER, M.; ROE, S. News from the classification & indexing section. *IFLA Metadata Newsletter*, v. 2, n. 1, june, 2016. Disponível em: <http://www.ifla.org/files/assets/classification-and-indexing/metadata_newsletter-20160803.pdf>. Acesso em: 25 set. 2016.

MODELO DE LEITURA DOCUMENTÁRIA PARA INDEXAÇÃO DE LITERATURA INFANTIL DO GÊNERO FÁBULA: ESQUEMAS TEXTUAIS PARA FUNDAMENTAÇÃO DE ESTRATÉGIAS METACOGNITIVAS

Roberta Caroline Vesu Alves

1 Introdução

A análise de assunto e identificação de conceitos em indexação, para fins de representação e recuperação da informação consiste em um processo cognitivo, que demanda análise e síntese. Também, a indexação é constituída de etapas que se iniciam pela leitura documentária (SILVA; FUJITA, 2004). Conforme verificado em Fujita (2013), a quantidade de etapas para indexação pode variar, dependendo do autor que a descreve e do que se observa durante a indexação. Mas, independentemente da quantidade de etapas, a indexação tem por finalidade obter uma representação de conceitos, para fins de recuperação da informação do documento. Ainda segundo a autora,

> A representação do conhecimento contido em conteúdos documentários é o objetivo principal da indexação, entendido de modo mais amplo. Em âmbito mais específico e considerando-se as perspectivas teóricas da indexação conceitual, a representação por conceitos assume função preponderante entre o significado do conteúdo documentário e o termo que o representa. (FUJITA, 2013, p. 44).

Nesse sentido, o termo utilizado para representação tem caráter de significante e está relacionado com o conceito, que possui caráter de significado, ideia e unidade de conhecimento. A representação ocorre também mediante o referente, o contexto dado por conteúdos documentários e a linguagem documentária para indexação e recuperação da informação (CAMPOS, 2001; FIORIN, 2002; SAUSSURE, 1999).

A Análise Documentária, de acordo com os estudos de Silva e Fujita (2004), é proveniente da corrente teórica francesa de J.-C. Gardin (1925-2013), também denominada Análise Documental, e compreende inicialmente a indexação como resultado da representação por meio de linguagens documentárias. Além disso, apresenta diferenciação das pesquisas espanholas, que a compreende enquanto o tratamento de forma e conteúdo, este relacionado com o "[...] tratamento temático da informação e destina-se à representação condensada do assunto intrínseco ou extrínseco tratado em um determinado documento" (SILVA; FUJITA, 2004, p. 137), entre outros aspectos.

Contudo, o conceito de indexação surgiu com a elaboração de índices, mas passou a significar a análise e identificação de assuntos do documento, além da representação (SILVA; FUJITA, 2004). A indexação enquanto análise de assuntos e identificação de conceitos em conteúdo documentário, em Análise Documentária, foi desenvolvida de modo interdisciplinar também com apoio teórico da Linguística Textual e Psicologia Cognitiva, principalmente, por pesquisadores espanhóis, por exemplo, Pinto e Gálvez (1999) e Pinto Molina (1993), e brasileiros como Fujita (1999, 2003, 2004), Neves (2004), Fujita e Rubi (2006b), Neves, Dias e Pinheiro (2006), Guimarães, Moraes e Guarido (2007), Alves (2008) e Alves et al. (2016).

O Modelo de Leitura Documentária para o texto científico compreende os questionamentos para a identificação de conceitos, isso como recurso estratégico de inferência ao texto, considerando a super estrutura e macro estrutura textual (FUJITA; RUBI, 2006b).

Diante disso, objetivou-se propor um Modelo de Leitura Documentária para identificação de conceitos ficcionais da literatura infantil, especificamente do gênero discursivo fábula infantil, embasado em

esquemas textuais e estratégias metacognitivas. Esses aspectos demandam o entendimento e explicitação da estrutura textual e das estratégias metacognitivas desejadas para fundamentar a leitura documentária. Para isso, utilizou-se a metodologia de estudo bibliográfico, com análise qualitativa das informações da literatura científica para verificação e proposta de esquemas textuais e estratégias metacognitivas, bem como a aplicação dessa proposta de metodologia em amostra aleatória de fábula infantil.

O Modelo de Leitura Documentária de Fujita e Rubi (2006b) e Fujita (2009) foi considerado como base para o desenvolvimento da proposta do novo modelo voltado ao texto de fábula infantil, como também as estruturas textuais estabelecidas por Van Dijk (1996, 2000, 2004), os processamentos mentais explicados por Neves (2011), além das teorias que explicam as categorias conceituais ficcionais, segundo aspectos teóricos da Teoria da Narrativa (D'ONOFRIO, 2007; GANCHO, 2014).

Verificou-se por meio do desenvolvimento da proposta do Modelo de Leitura Documentária e aplicação para exemplificação da indexação, que o texto do gênero fábula infantil contém categorias conceituais embasadas em sua estrutura textual. Além disso, constatou-se que elementos importantes para a recuperação podem ser indexados a partir de aspectos característicos do gênero textual e da narrativa, que se referem ao gênero da fábula infantil, tipo de narrador, aspectos de personagens, espaço, tempo, enredo e moral da história. A proposta de modelo de leitura foi estabelecida com base em esquemas textuais, que podem ser utilizados como estratégias metacognitivas durante a leitura documentária, com o intuito de facilitar a inferência e identificação de conceitos durante a indexação.

2 O modelo de leitura documentária para indexação de texto científico

A leitura documentária para identificação de conceitos em texto científico pode ser realizada com base em estratégias de exploração de estruturas textuais e de abordagem sistemática para identificação desses conceitos, conforme o Modelo de Leitura Documentária. Esse modelo teve seu método desenvolvido por Fujita (2003b) e Fujita e Rubi (2006a, 2006b), tendo como apoio a metodologia de Protocolo Verbal, conforme também Fujita,

Nardi e Santos (1998), Fujita, Nardi e Fagundes (2003), entre outros pesquisadores, além do aprimoramento do modelo por Fujita (2009).

Além disso, verificou-se que o Modelo de Leitura Documentária tem como base os estudos da pesquisadora Maria de Fátima Gonçalves Moreira Tálamo, a Norma 12.676 da Associação Brasileira de Normas Técnicas (ABNT) e estudos sobre o sistema *PREserved Context Indexing System* (PRECIS), segundo Fujita e Rubi (2006a, 2006b).

De modo geral, a abordagem sistemática para identificação de conceitos buscou responder questionamentos sobre as categorias temáticas: o que? (categoria essencial), quando?, onde?, como? (categorias acessórias). Pois foram consideradas como elementos fundamentais para os modelos de leitura de indexação (FUJITA; RUBI, 2006b).

O Modelo de Leitura Documentária, segundo Fujita e Rubi (2006b), consiste em "[...] uma metodologia baseada no uso de estratégias de leitura". Essa metodologia para indexação, ainda segundo as autoras, considera para identificação de conceitos, tanto o processo de cognição de leitura, como também o conhecimento prévio do leitor sobre aspectos de mundo, profissionais, linguísticos, textuais, além de conhecimentos específicos de indexadores especialistas, sendo todos esses aspectos relacionados com estratégias de leitura, esquemas mentais e estruturas textuais (FUJITA; RUBI, 2006b).

O estudo das autoras sobre o Modelo de Leitura Documentária foi aprimorado com observação da indexação por meio de Protocolo Verbal, em formação de bibliotecários indexadores de bibliotecas universitárias. Portanto, de acordo com Fujita e Rubi (2006b), as estratégias de leitura dos indexadores que fundamentaram o Modelo de Leitura Documentária são, principalmente, os aspectos de:

a) estratégias metacognitivas;

b) conhecimento prévio linguístico, textual e de mundo (leitor inato sob o ponto vista linguístico e cognitivo);

c) domínio da estrutura textual;

d) exploração da estrutura textual como estratégia de leitura documentária para identificação e seleção de conceitos, durante a análise de assunto;

e) dependência da estratégia de exploração da estrutura textual na identificação de conceitos.

Vale destacar que outro aspecto importante para a metodologia de Modelo de Leitura Documentária é a abordagem sistemática para identificação de conceitos, que ocorre por meio de categorias de questionamentos ao texto e análise dessas categorias temáticas, com relação às partes do texto, ou seja, relacionando-as com a estrutura textual (FUJITA; RUBI, 2006b).

A estrutura textual, ou partes do texto científico, a ser explorada em leitura documentária, ainda segundo as autoras, é prevista no Modelo de Leitura Documentária e consiste em: Título em português; Título em inglês; Autoria; Resumo do trabalho científico; Palavras-chave; *Abstract*; *Keywords*; Introdução; Materiais e métodos; Resultados; Figuras; Discussão dos resultados; Conclusões; e Referências bibliográficas.

Posteriormente, para a compreensão global do texto e ização do conteúdo pertinente a cada parte, é necessário observar se as seguintes partes do texto contêm aspectos relacionados com:

> **Introdução**: explicação do assunto principal com referencial teórico, contendo os objetivos com o tema principal do trabalho ao final da introdução;
>
> **Materiais e métodos**: descrição de materiais e métodos utilizados, processos, técnicas, amostragem;
>
> **Resultados**: compatibilidade com objetivos enunciados, materiais e métodos utilizados, com o uso, às vezes de figuras, gráficos, tabelas, fotografias, etc;
>
> **Discussão dos resultados**: verificação dos resultados a partir do referencial teórico utilizado;
>
> **Conclusões**: verificação dos objetivos propostos;
>
> **Referências bibliográficas**. (FUJITA; RUBI, 2006b, grifo nosso).

Desse modo, é possível compreender o conteúdo como um todo e começar a explorar a estrutura textual para identificar os conceitos de cada parte importante, que são revelados por meio de questionamento para indexação. De acordo com Fujita e Rubi (2006b), "conceito é a formulação de uma ideia por palavras", sendo que o conceito pode ser representado por uma palavra do texto, também os conceitos essenciais do documento são correspondentes ao seguinte entendimento dos componentes do texto científico:

> * **OBJETO**: é algo ou alguém que está sob estudo do pesquisador.
>
> * **AÇÃO**: processo sofrido por algo ou alguém.
>
> * **AGENTE**: aquele ou algo que realizou a ação.
>
> * **MÉTODOS**: métodos utilizados para realização da pesquisa.
>
> * **FÍSICO OU AMBIÊNCIA**: físico onde foi realizada a pesquisa.
>
> * **CAUSA E EFEITO**:
>
> **- causa** => razão ou motivo. Aquilo ou aquele que faz com que uma coisa exista ou aconteça (antecedente); está vinculada à identificação da AÇÃO.
>
> **- efeito** => produto de uma causa. Resultado de um ato qualquer (consequência); está vinculado ao resultado da AÇÃO realizada.
>
> Assim, o suposto efeito ou consequente deve variar cada vez que faz variar a suposta causa ou antecedente. (FUJITA; RUBI, 2006b, grifo nosso).

Esses componentes do texto científico – que são objeto, ação, agente, métodos, físico ou ambiência, causa e efeito –, formam categorias conceituais, que contribuem para a identificação dos conceitos correspondentes a cada uma durante a leitura documentária, considerando para isso a estrutura textual do texto científico. A estrutura textual mais relevante do texto científico compreende: introdução (contém o objetivo com o tema mais importante do texto), metodologia, resultados e conclusão. A seguir, o quadro 1 demonstra as partes do texto, o conteúdo de cada parte textual e as categorias conceituais.

Quadro 1– Identificação de conceitos mediante exploração da estrutura textual.

PARTES DO TEXTO	CONTEÚDO PERTINENTE	CONCEITOS DE
Introdução (objetivos)	Referencial teórico Tema: objetivos* Objetivos	OBJETO AGENTE AÇÃO
Metodologia	Descrição de materiais, métodos, processos e técnicas utilizados.	MÉTODOS FÍSICO MATERIAIS
Resultados Discussão dos resultados	Compatibilidade com objetivos enunciados e materiais e métodos utilizados, mostrados, às vezes em tabelas; Verificação dos resultados a partir do referencial teórico utilizados	CAUSA E EFEITO

Fonte: Fujita e Rubi (2006b).

Diante disso, vale destacar que o resultado da indexação poderá ser obtido mais facilmente com a utilização do questionamento ao texto, mesmo que nem todas as questões possam ser respondidas, isso conforme as questões estabelecidas por Fujita e Rubi (2006b):

1. O assunto contém uma ação (podendo significar uma operação, um processo etc.)?

2. O documento possui em seu contexto um objeto sob efeito desta ação?

2.1 O objeto identificado pode ser considerado como parte de uma totalidade?

2.2 O objeto identificado possui características ou atributos particulares?

3. O documento possui um agente que praticou esta ação?

4. Para estudo do objeto ou implementação da ação, o documento cita e/ou descreve modos específicos, por exemplo: instrumentos especiais, técnicas, métodos, materiais e equipamentos?

5. A ação, objeto e agente são considerados no contexto de um lugar específico ou ambiente?

6. Considerando que a ação e o objeto identificam uma causa, qual é o efeito desta causa? Causa [...] (ação+objeto).

Os questionamentos também contribuem para a compreensão do texto e identificação de conceitos, mas a identificação de conceitos pode ser entendida quando os conceitos são comparados com exploração da estrutura textual, conforme estabelecido pelas autoras no quadro 2, a seguir:

Quadro 2– Modelo de Leitura Documentária para textos científicos: identificação de conceitos por questionamento em partes da estrutura textual.

CONCEITO (ANÁLISE CONCEITUAL)	QUESTIONAMENTO (NORMA 12676)	PARTE DA ESTRUTURA TEXTUAL
OBJETO	O documento possui em seu contexto um objeto sob efeito de uma atividade?	INTRODUÇÃO (OBJETIVOS)
AÇÃO	O Assunto contém um conceito ativo (por exemplo, uma ação, uma operação, um processo etc?)	INTRODUÇÃO (OBJETIVOS)
AGENTE	O documento possui um agente que praticou uma ação?	INTRODUÇÃO (OBJETIVOS)
MÉTODOS DO AGENTE	Este agente refere-se a modos específicos pra realizar a ação (por exemplo, instrumento especiais, técnicas ou métodos)	METODOLOGIA
LOCAL OU AMBIÊNCIA	Todos estes fatores são considerados no contexto de um lugar específico ou ambiente?	METODOLOGIA
CAUSA E EFEITO	São identificados algumas variáveis dependentes ou independentes?	RESULTADOS; DISCUSSÃO DE RESULTADOS
PONTO DE VISTA DO AUTOR; PERSPECTIVA	O assunto foi considerado de um ponto de vista, normalmente não associado com o campo de estudo (por exemplo, um estudo sociológico ou religioso)?	CONCLUSÕES

Fonte: Fujita e Rubi (2006b).

O método demonstrado para o Modelo de Leitura Documentária de Fujita e Rubi (2006b) permite a identificação de conceitos para indexação de modo estratégico, pois considera as estruturas textuais do texto científico e consciente, portanto, estabelece caminhos para o reconhecimento de esquemas textuais e uso de estratégias metacognitivas de leitura documentária.

Verificou-se que a falta de procedimentos sistemáticos para a identificação de conceitos reforçou a necessidade de desenvolvimento da metodologia de análise de assuntos para a leitura documentária, buscando diminuir as dificuldades de identificação de conceitos (FUJITA; RUBI, 2006b).

Destaca-se que, a partir disso, o aprimoramento de Fujita (2009) sobre o Modelo de Leitura Documentária, estabelecendo-o conforme quadro 3, que foi adaptado do original da autora, pois foi retirado para este estudo a última coluna com os exemplos de conceitos identificados.

Quadro 3–Versão adaptada do Modelo de Leitura Documentária.

CONCEITO	QUESTIONAMENTO PARA IDENTIFICAÇÃO DE CONCEITOS	PARTES DA ESTRUTURA TEXTUAL
OBJETO e PARTE(S) DO OBEJTO (algo ou alguém que está sob estudo do pesquisador)	O documento possui em seu contexto um objeto sob efeito desta ação?	INTRODUÇÃO (OBJETIVOS)
AÇÂO (processo sofrido por algo ou alguém)	O assunto contém uma ação (podendo significar uma operação, um processo etc.?)	INTRODUÇÃO (OBJETIVO)
AGENTE (aquele ou algo que realizou a ação)	O documento possui um agente que praticou esta ação?	INTRODUÇÃO (OBJETIVO)
MÉTODO (métodos utilizados para relaização da pesquisa)	Para estudo do objeto ou implementação da ação, o documento cita e/ou descreve modos específicos, por exemplo: instrumentos especiais, técnicas, métodos, materiais e equipamentos?	METODOLOGIA
LOCAL OU AMBIÊNCIA (físico onde foi realizada a pesquisa)	Todos estes fatores são considerados no contexto de um lugar específico ou ambiente?	METODOLOGIA
CAUSA E EFEITO Causa (ação+objeto)/ Efeito	Considerando que a ação e objeto identificam uma causa, qual é o efeito desta causa?	RESULTADOS; DISCUSSÃO DE RESULTADOS; CONCLUSÕES

Fonte: Adaptado de Fujita (2009, p. 437).

A adaptação de Fujita (2009) para o Modelo de Leitura Documentária, com fins de facilitar o ensino de indexação, permitiu mais visibilidade e entendimento sobre em que consistem as categorias conceituais, especialmente sobre métodos, bem como outros aspectos que demonstram a mudança no modelo, conforme destacado pela autora a seguir:

- Acima do quadro, recomenda-se incluir uma síntese dos procedimentos descritos no manual explicativo como um "passo-a-passo";
- A identificação dos conceitos ação, objeto e agente merece maior atenção na adequação do modelo com a finalidade de esclarecer e exemplificar a identificação desses conceitos em vários aspectos que se recomenda estarem descritos como observações estratégicas;
- alteração da denominação do conceito "Métodos do agente" para "Métodos" e a exclusão do conceito "Ponto de vista do autor; perspectiva", uma vez que raramente é utilizado;
- incluir na primeira coluna breves explicações sobre cada conceito de forma a orientar o indexador no uso específico de cada um;
- incluir na coluna de "Questionamento" explicação sobre o objetivo do questionamento;
- Incluir quarta coluna com indicação de termos que representam os conceitos para exemplificar o processo de identificação de conceitos para análise de assunto;

A utilização e aprimoramento do Modelo de Leitura Documentária permitiram, conforme Fujita (2009), um maior entendimento sobre os conceitos a serem identificados no texto e sobre onde encontrá-los, por exemplo, os conceitos de objeto, ação e agente estão interligados, são principais e devem ser identificados nesta ordem, entre outros aspectos. Além disso, de acordo com a mesma autora, o modelo facilita "[...] o processo de indexação, além de mostrar-se uma ferramenta de auxílio, pois orienta os indexadores aprendizes sobre o que procurar e onde encontrar" (FUJITA, 2009, p. 441).

Outro aspecto importante que contribuiu para o aprimoramento do Modelo de Leitura Documentária, segundo Fujita (2009), foi o uso da metodologia de Protocolo Verbal no ensino de leitura documentária para indexação, que propiciou a revelação de estratégias e dificuldades a serem equacionadas ou aprimoradas. Isso porque a metodologia do Protocolo Verbal promove, entre outros aspectos, a coleta de dados introspectivos durante a realização de uma tarefa, desvendando os processos cognitivos como os esquemas e estratégias metacognitivas (FUJITA, 2009).

3 A literatura infantil e gênero discursivo fábula infantil

A literatura infanto-juvenil, em geral, tem alcançado atualmente posição de relevância no Brasil, contradizendo, segundo Souza (2006, p. 17), a "[...] desvalorização e marginalidade a que essa produção literária ficou sujeita durante muito tempo, tendo sido sempre considerada uma literatura menor". Essa desvalorização ocorreu não só por se tratar de um texto voltado ao leitor infantil, mas também por ter sido tradicionalmente utilizado para fins morais e pedagógicos até meados da década de 1970, quando os novos autores inovaram a produção cultural para crianças e jovens, buscando estimular o lazer e elementos artísticos (SOUZA, 2006).

Segundo Cademartori (2010), a literatura infantil é caracterizada, entre outros aspectos, pela forma como é direcionada para a idade do leitor, ou seja, para a competência de leitura estimada para cada faixa etária e se tornou inseparável do universo educacional. Além disso, conforme a autora,

> As obras infantis que respeitam seu público são aquelas cujos textos têm potencial para permitir ao leitor infantil possibilidade ampla de atribuição de sentidos àquilo que lê. A literatura infantil digna do nome estimula a criança a viver uma aventura com a linguagem e seus efeitos, em lugar de deixá-la cerceada pelas intenções do autor, em livros usados em transportes de intenções diversas, entre elas o que se passou a chamar de "politicamente correto", a nova face do interesse pedagógico, que quer se sobrepor ao literário. (CADEMARTORI, 2010, p. 17).

Nesse contexto, segundo afirma a mesma autora, a importância das narrativas clássicas para crianças é amplamente conhecida, além disso,

os conflitos vivenciados por personagens ficcionais desse tipo de texto permitem aos ouvintes e leitores a identificação, reflexão e atribuição de novas perspectivas para esses conflitos (CADEMARTORI, 2010). De acordo com Souza (2006, p. 53), a importância da literatura infantil também está em estimular a imaginação e o gosto pela leitura, e ainda,

> A literatura infanto-juvenil é a primeira forma escrita de contato com a criança e do jovem com as tradições culturais e literárias de seu povo. Ao mesmo tempo que promove recreação também cultiva valores necessários à vida em sociedade e favorece o raciocínio e a inteligência da criança e do jovem.

Historicamente, as primeiras literaturas para crianças, além das religiosas, eram as histórias populares e de tradição oral, mas também agradavam pessoas de todas as idades. As fábulas, por exemplo, existem desde a Antiguidade (SOUZA, 2006).

O texto da fábula, segundo Souza (2006), apresenta características de:

a) narração e texto curtos;
b) acontecimentos fictícios relacionados com a vida humana;
c) dupla finalidade de instruir e divertir;
d) moralidade verossímil;
e) personagens animais que falam e demonstram atos e conflitos;
f) animais não são escolhidos ao acaso, pois suas características são aproveitadas;
g) animais também mostram vícios humanos;
h) críticas ao caráter, costumes e às maldades humanas.

Segundo Duarte (2013), as fábulas são famosas por conter em seu final a moral da história, mas essa moral pode ser apresentada de modo explícito ou implícito no texto. Além disso, o gênero discursivo fábula infantil pode ser apresentado em texto rimado (poesia e poema) ou em prosa. Por isso, para este estudo foram consideradas as fábulas de Esopo

apresentadas com narrativas em prosa. Esopo, segundo Duarte (2013) e Tufano e Nóbrega (S.d.), foi um escravo muito famoso, autor e difusor de diversas fábulas, que viveu na Grécia a partir do final do século V a.C., mas os detalhes de sua vida ainda motivam divergências entre os historiadores.

As características da fábula infantil também são constituídas das características do texto ficcional de narrativa em prosa, que compreendem, segundo D'Onofrio (2007), Gancho (2014) e Van Dijk (1996, 2000, 2004), em:

a) **estruturas textuais**: microestrutura (palavras da superfície textual), macroestrutura (elementos semânticos) e superestrutura (elementos de organização do texto);

b) **personagens:** principais (protagonista e antagonista, com suas características marcantes) e secundários;

c) **espaço ou** : cenário ou em que se passa a ação ou enredo; ambiente psicológico, social, moral, econômico etc. (por exemplo, hostil, pobre, agradável, requintado etc.);

d) **enredo**: fatos, ações e acontecimentos vivenciados por personagens mediante espaço, tempo, contados por narrador, apresentando como parte central o conflito (enredo composto também de apresentação, complicação, clímax e desfecho);

e) **tempo**: época da história (exemplo: Idade Média e inverno); duração da história (enredo curto ou enredo de muitos anos); tempo cronológico (mensurável em horas, dias, meses, anos etc.); e tempo psicológico (ordem determinada pela imaginação do narrador ou do personagem);

f) **tipos de narrador**: narrador personagem (verbos em 1ª pessoa), pode ser protagonista ou testemunha; narrador onisciente e observador (verbos em 3ª pessoa).

Além disso, os gêneros discursivos em conjunto com essas características de narrativa em prosa são elementos requisitados por usuários de biblioteca escolar, como crianças, adolescentes e professores, conforme pôde se observar em atuação profissional. São, pois, categorias conceituais

que expressam o que os usuários gostariam de ler, por exemplo: gêneros de aventura, romance, fábula, poesia; temáticas que demonstram superação de preconceito, comportamento adequado, luta entre o bem e o mal; personagens bruxos e animais favoritos; conflito entre amizades; tempo da narrativa e tipo de narrador (conforme leitura indicada pelo professor); etc.

Essas características textuais, portanto, também expressam categorias conceituais importantes para a indexação, propiciando, posteriormente à indexação, os meios de recuperação da informação, respeitando também as necessidades de informação apresentadas pelos usuários observados.

Sabe-se que ainda muito deve ser feito no Brasil para que crianças e adolescentes tomem contato e desfrutem da literatura ficcional e da leitura, de modo geral. Mas, o que não deve ocorrer em bibliotecas públicas e escolares, é que o leitor deixe de encontrar a literatura de acordo com seu gosto, ou com o que gostaria de ler no momento. Para amenizar essa a falta de meios de indexação e recuperação da informação sobre a literatura infantil, foi desenvolvida uma proposta de Modelo de Leitura Documentária para fins indexação do gênero fábula infantil, visando também contribuir com os estudos sobre indexação.

4 Proposta de modelo de leitura documentária de literatura do gênero fábula infantil: esquemas textuais para fundamentação de estratégias metacognitivas

Os esquemas textuais ou superestruturas, segundo Neves (2011, p. 36), consistem em um "[...] conjunto de conhecimentos adquiridos à proporção que lemos diversos tipos de textos e efetuamos correlação entre eles". Nesse sentido, a superestrutura, de acordo com Van Dijk (2000, 2004), estabelece a forma e a organização textual, caracterizando o tipo de texto em esquemas formais, contribuindo para expressar determinado sentido em âmbito macroestrutural. Outros aspectos da estrutura textual são, por conseguinte, a macroestrutura, que consiste, entre outros aspectos, em conteúdo semântico parcial e global, como também a microestrutura em nível superficial, com proposições básicas (VAN DIJK, 1996, 2000, 2004).

O conhecimento da organização da estrutura textual e os conhecimentos prévios permitem ao leitor indexador interagir com o texto por

meio de estratégias metacognitivas, ou seja, direcionando a leitura documentária para as partes específicas do texto, ao mesmo tempo em que direciona para essas partes a compreensão do texto necessária aos fins de indexação e recuperação da informação (FUJITA, 1999).

Segundo Neves (2011), a cognição denota tipos específicos de representação, que pode ser, por exemplo, sobre objetos, fatos e informações provenientes do meio. Mas, a metacognição significa "[...] o conhecimento sobre o próprio conhecimento, avaliação, regulação e organização dos processos cognitivos" (NEVES, 2011, p. 29). Nesse contexto, a estratégia "[...] refere-se às informações sobre meios e processos ou ações que permitem ao sujeito atingir objetivos com maior eficácia em determinada tarefa" (NEVES, 2011, p. 33).

Portanto, a estratégia utilizada para atingir objetivos de leitura documentária e indexação é a observação e compreensão de esquemas textuais (superestruturas que apontam para o conteúdo textual), em conjunto com a metacognição, que gerencia de modo explícito os conhecimentos e organiza os processos cognitivos.

Verificou-se que o Modelo de Leitura Documentária é fundamentado por procedimentos que consideram os esquemas textuais (domínio da estrutura textual), as estratégias metacognitivas e abordagem sistemática, esta, por meio de questionamentos, entre outros aspectos. Por isso, as etapas para construção do Modelo de Leitura Documentária de Literatura do Gênero Fábula Infantil consistem em:

a) identificação das estruturas textuais;

b) conceitos e questionamentos relacionados com as partes textuais;

c) exemplificações da aplicação do Modelo de Leitura Documentária de Literatura do Gênero Fábula Infantil em fábulas selecionadas aleatoriamente.

Para a identificação das estruturas textuais do gênero fábula infantil, foi considerado, portanto, o tipo de gênero discursivo, a microestrutura, macroestrutura e superestrutura, bem como as partes do texto narrativo, que também o caracterizem. Esses elementos textuais são impor-

tantes para descrever o conteúdo do texto, ao mesmo tempo em que são importantes para recuperação da informação aos usuários a que esse gênero textual se destina.

O quadro 4, a seguir, reúne as estruturas textuais e o conteúdo esperado para cada parte e característica desse tipo de texto.

Quadro 4 – Estruturas textuais e conteúdos pertinentes.

PARTES DO TEXTO NARRATIVO	CONTEÚDO PERTINENTE
ESTRUTURAS TEXTUAIS GERAIS	Conteúdo semântico em macroestrutura, revelados pela microestrutura; Superestrutura conforme as partes do texto narrativo e gênero discursivo.
GÊNERO DISCURSIVO	Fábula infantil.
TÍTULO DA FÁBULA INFANTIL	Geralmente revela os personagens principais.
ENREDO	Toda a história (apresentação, complicação, clímax e desfecho), principalmente o conflito. Portanto, é importante desmembrar todas as partes do enredo para indexação.
ENREDO – APRESENTAÇÃO: PERSONAGENS	Personagens principais (protagonista/bem e antagonista/mal) e suas características marcantes.
ENREDO – APRESENTAÇÃO: ESPAÇO OU	Cenário: descrição do ; Ambiente: descrição subjetiva do ambiente.
ENREDO – APRESENTAÇÃO: TEMPO	Época da história; Duração da história; Tempo cronológico ou psicológico.
ENREDO – COMPLICAÇÃO	Após uma normalidade é iniciado um conflito.
ENREDO – CLÍMAX	Ponto máximo do conflito, antes de seu desfecho.
ENREDO – DESFECHO	Desfecho ou conclusão do conflito, estabelecendo outra normalidade.
MORAL DA HISTÓRIA (NO FINAL DO TEXTO)	Ensinamento filosófico e/ou moral alertando para o bem comum.
NARRADOR	Narrador personagem (protagonista ou testemunha); Narrador onisciente ou observador.

Fonte: Elaborado pela autora, com base em D'Onofrio (2007), Fujita e Rubi (2006b), Fujita (2009), Gancho (2014) e Van Dijk (1996, 2000, 2004).

A partir do estabelecimento da identificação das estruturas textuais, passou-se a propor os conceitos e questionamentos relacionados com cada parte, para o Modelo de Leitura Documentária para o Gênero Fábula Infantil, conforme quadro 5.

Quadro 5 – Modelo de Leitura Documentária para o Gênero Fábula Infantil: conceitos e questionamentos direcionados para partes textuais.

PARTES DO TEXTO NARRATIVO	QUESTIONAMENTOS	CONCEITOS
GÊNERO DISCURSIVO	Qual o gênero e subgênero discursivo?	**GÊNERO DISCURSIVO**
ENREDO – APRESENTAÇÃO: PERSONAGENS	Quem são os personagens principais? Quais as características importantes dos personagens principais?	**PERSONAGENS**
ENREDO – APRESENTAÇÃO: ESPAÇO OU LOCAL	Em que cenário ocorrea ação ou enredo? Em que ambiente ocorre a ação ou enredo?	**ESPAÇO OU LOCAL**
ENREDO – APRESENTAÇÃO: TEMPO	Em que época se passa a história? Quanto tempo dura a história? O tempo é apresentado de modo cronológico ou psicológico?	**TEMPO**
ENREDO – COMPLICAÇÃO	Qual o conflito apresentado?	**CONFLITO**
ENREDO – CLÍMAX	Qual o clímax ou ponto máximo do conflito?	**CLIMAX**
ENREDO – DESFECHO	Qual o desfecho do conflito?	**DESFECHO**
MORAL DA HISTÓRIA (NO FINAL DO TEXTO)	Qual a moral da história?	**MORAL DA HISTÓRIA**
NARRADOR	Qual o tipo de narrador?	**NARRADOR**

Fonte: Elaborado pela autora, com base em D'Onofrio (2007), Fujita e Rubi (2006b), Fujita (2009), Gancho (2014) e Van Dijk (1996, 2000, 2004).

O Modelo de Leitura Documentária para o Gênero Fábula Infantil proposto acima reflete os tipos de conceitos importantes para representação e recuperação da informação, bem como a abordagem sistemática de questionamento ao texto e a relação desses aspectos com a estrutura textual pertinente ao tipo de texto escolhido para esta proposta.

Diante disso, o modelo foi aplicado em amostra aleatória, conforme quadros 6 e 7, exemplificando como pode ocorrer a indexação para identificação de conceitos em duas fábulas infantis do autor Esopo (2013).

Quadro 6 – Exemplificação da aplicação do modelo para indexação da fábula infantil "A cigarra e as formigas" de Esopo.

CONCEITOS	QUESTIONAMENTOS	TERMOS IDENTIFICADOS
GÊNERO DISCURSIVO	Qual o gênero e subgênero discursivo?	**Fábula. Fábula infantil.**
PERSONAGENS	Quem são os personagens principais?	**Cigarra. Formigas.**
	Quais as características importantes dos personagens principais?	**Cigarra cantora.** **Formigas trabalhadoras.**
ESPAÇO OU LOCAL	Em que cenário ocorre a ação ou enredo?	**[Não revela].**
	Em que ambiente ocorre a ação ou enredo?	**[Não revela].**
TEMPO	Em que época se passa a história?	**Inverno.**
	Quanto tempo dura a história?	**Narrativa em prosa curta.**
	O tempo é apresentado de modo cronológico ou psicológico?	**Tempo cronológico. Tempo psicológico.**
CONFLITO	Qual o conflito apresentado?	**Fome. Pedido de alimento.**
CLIMAX	Qual o clímax ou ponto máximo do conflito?	**Não trabalhou. Cantou.**
DESFECHO	Qual o desfecho do conflito?	**Risos. [As formigas avisaram a cigarra]: Dance no inverno.**
MORAL DA HISTÓRIA	Qual a moral da história?	**Não descuidar das tarefas.**
NARRADOR	Qual o tipo de narrador?	**Narrador onisciente.**

Fonte: Elaborado pela autora, com base em D'Onofrio (2007), Esopo (2013, p. 154), Fujita e Rubi (2006b), Fujita (2009), Gancho (2014).

A segunda exemplificação da aplicação do modelo para identificação de conceitos ocorreu conforme quadro 7.

Quadro 7– Exemplificação da aplicação do modelo para indexação da fábula infantil "A tartaruga e a lebre" de Esopo.

CONCEITOS	QUESTIONAMENTOS	TERMOS IDENTIFICADOS
GÊNERO DISCURSIVO	Qual o gênero e subgênero discursivo?	**Fábula.Fábula infantil.**
PERSONAGENS	Quem são os personagens principais?	**Tartaruga. Lebre.**
	Quais as características importantes dos personagens principais?	**Tartaruga lenta. Tartaruga persistente.** **Lebre veloz. Lebre displicente.**
ESPAÇO OU LOCAL	Em que cenário ocorre a ação ou enredo?	**Um percurso.**
	Em que ambiente ocorre a ação ou enredo?	**[Não revela].**
TEMPO	Em que época se passa a história?	**[Não revela].**
	Quanto tempo dura a história?	**Narrativa em prosa curta.**
	O tempo é apresentado de modo cronológico ou psicológico?	**Tempo cronológico.**
CONFLITO	Qual o conflito apresentado?	**Disputa de velocidade.**
CLIMAX	Qual o clímax ou ponto máximo do conflito?	**Lebre adormeceu.**
DESFECHO	Qual o desfecho do conflito?	**Tartaruga venceu.**
MORAL DA HISTÓRIA	Qual a moral da história?	**O esforço vence a natureza relapsa.**
NARRADOR	Qual o tipo de narrador?	**Narrador onisciente.**

Fonte: Elaborado pela autora, com base em D'Onofrio (2007), Esopo (2013, p. 496), Fujita e Rubi (2006b), Fujita (2009), Gancho (2014).

A aplicação do Modelo de Leitura Documentária para o Gênero Fábula Infantil demonstrou a identificação de conceitos importantes para representação e recuperação, condizentes com o conteúdo, termos e características das fábulas utilizadas para indexação. Contudo, não é possível nesta pesquisa traçar um paralelo ou análise entre os termos identificados e os termos que possivelmente seriam selecionados mediante a tradução com o uso de linguagem documentária. Pois esta proposta de modelo de

leitura se limita ao estabelecimento da identificação de estruturas textuais e desenvolvimento de questionamentos para identificação de conceitos e termos do texto.

5 Considerações finais

A proposta de Modelo de Leitura Documentária para indexação de fábula foi embasada em esquemas textuais para fundamentação de estratégias metacognitivas de leitura, proporcionando um meio para identificação de conceitos desse tipo de literatura infantil. Também teve como base o modelo de leitura para texto científico, no que se refere aos parâmetros de utilização de estrutura textual como estratégia metacognitiva para identificação de conceitos, e abordagem sistemática de questionamento ao texto.

A exemplificação da aplicação da proposta do modelo de leitura elaborado demonstrou que os conceitos e termos puderam ser identificados nos textos escolhidos em amostra, mas, o modelo ainda precisa ser testado por indexadores, preferencialmente, com pesquisas que comprovem o resultado do uso desse modelo por meio de Protocolo Verbal. Ainda, sugere-se que outros estudos sejam realizados para demonstrar a aprendizagem do modelo proposto, bem como seu aprimoramento mediante fundamentação de novas necessidades de representação e recuperação.

As dificuldades de tratamento e indexação da literatura ficcional em bibliotecas escolares e públicas ainda são grandes no Brasil. Entretanto, é imprescindível que pesquisas e ações profissionais promovam o desenvolvimento de procedimentos de indexação, como também as linguagens documentárias voltadas para literatura ficcional, contribuindo desse modo para o avanço da representação e recuperação desse tipo de informação, que é muito importante para seus usuários.

Referências

ALVES, R. C. V. *Análise Documental de Textos Literários Infanto-Juvenis*: perspectivas metodológicas com vistas à identificação do tema. 165 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, UNESP, Marília, 2008.

ALVES, R. C. V. et al. Estratégias metacognitivas para análise de assunto: aspectos teóricos de superestrutura e esquemas sobre textos literários infanto-juvenis. *Informação & Sociedade: Estudos*, João Pessoa, v. 26, n. 1, p. 07-19, jan./abr. 2016. Disponível em: <http://www.ies.ufpb.br/ojs/index.php/ies/article/view/21957/15525>. Acesso em: 25 nov. 2016.

CADEMARTORI, L. *O que é literatura infantil*. 2. ed. São Paulo: Brasiliense, 2010.

CAMPOS, M. L. A. *Linguagem Documentária*: teorias que fundamentam sua elaboração. Niterói: EdUFF, 2001.

D'ONOFRIO, S. *Forma e sentido do texto literário*. São Paulo: Ática, 2007.

DUARTE, A. Apresentação. In: ESOPO: fábulas completas. São Paulo: Cosac Naify, 2013.

ESOPO: fábulas completas. São Paulo: Cosac Naify, 2013.

FUJITA, M. S. L. A identificação de conceitos no processo de análise de assunto para indexação. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v. 1, n. 1, p. 60-90, jul./dez. 2003a. Disponível em: <http://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/2089/2219>. Acesso em: 08 dez. 2016.

FUJITA, M. S. L. A leitura do indexador: estudo de observação. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 4, n. 1, p. 101 - 116, jan./jun.1999. Disponível em: <http://www.brapci.inf.br/_repositorio/2010/11/pdf_78c2ce7393_0012667.pdf>. Acesso em: 20 out. 2015.

FUJITA, M. S. L. *A leitura documentária do indexador*: aspectos cognitivos e linguísticos influentes na formação do leitor profissional. 321f. Tese (Livre-Docência em Análise Documentária e Linguagens Documentárias Alfabéticas) – Faculdade de Filosofia e Ciências, UNESP, Marília, 2003b.

FUJITA, M. S. L. A Leitura Documentária na perspectiva de suas variáveis: leitor-texto-contexto. *DataGramaZero - Revista de Ciência da Informação*, v. 5, n. 4, ago. 2004. Disponível em: <http://www.datagramazero.org.br/ago04/Art_01.htm>. Acesso em: 20 out. 2015.

FUJITA, M. S. L. A representação documentária no processo de indexação com o modelo de leitura documentária para textos científicos e livros: uma abordagem cognitiva com protocolo verbal. *PontodeAcesso*, Salvador, v. 7, n. 1, p. 42-66, abr. 2013. Disponível em: <https://portalseer.ufba.br/index.php/revistaici/article/view/8135/5807>. Acesso em: 05 out. 2016.

FUJITA, M. S. L. Modelo de leitura documentária para indexação de textos científicos como metodologia de ensino sociocognitiva: análise da aplicabilidade com uso de protocolo verbal com vistas à sua adequação. In: BORGES, M. M.; CASADO, E. S. (Coord.). *A Ciência da Informação criadora de conhecimento*. Coimbra: Universidade de Coimbra, 2009. p. 431-448. Disponível em: <https://digitalis-dsp.uc.pt/jspui/bitstream/10316.2/31955/1/34-%20a%20ci%C3%AAncia%20da%20informa%C3%A7%C3%A3o%20criadora%20de%20conhecimento%20vol%20I.pdf?ln=pt-pt>. Acesso em: 08 dez. 2016.

FUJITA, M. S. L.; NARDI, M. I. A.; FAGUNDES, S. A. F. Observing documentar reading by verbal protocol. *Information Research*, v. 8, n. 4. July, 2003. Disponível em: <http://www.informationr.net/ir/8-4/paper155.html>. Acesso em: 08 dez. 2016.

FUJITA, M. S. L.; NARDI, M. I. A.; SANTOS, S. A leitura em análise documentária. *Transinformação*, Campinas, v. 10, n. 3, p. 13-3. set./dez., 1998.

FUJITA, M. S. L.; RUBI, M. P. Modelo de lectura profesional para la indización. *Scire*, Zaragoza, v. 12, n. 1, p. 47-70. 2006a. Disponível em: <http://ibersid.eu/ojs/index.php/scire/article/view/1585/1557>. Acesso em: 08 dez. 2016.

FUJITA, M. S. L.; RUBI, M. P. Um modelo de leitura documentária para a indexação de artigos científicos: princípios de elaboração e uso para a formação de indexadores. *DataGramaZero - Revista de Ciência da Informação*, v. 7, n. 3, jun. 2006b. Disponível em: <http://www.dgz.org.br/jun06/Art_04.htm>. Acesso em: 20 out. 2015.

GANCHO, C. V. *Como analisar narrativas*. 9. ed. São Paulo: Ática, 2014.

GUIMARÃES, J. A. C.; MORAES, J. B. E.; GUARIDO, M. D. M. Análisis documental de contenido de textos narrativos: bases epistemológicas y perspectivas metodológicas. *Ibersid*, Zaragoza, v. 1, p. 93-100, 2007. Disponível em: <http://www.ibersid.eu/ojs/index.php/ibersid/article/view/3267/3028>. Acesso em: 17 jan. 2008.

MOREIRO GONZÁLEZ, J. A. *El contenido de los documentos textuales:* su análisis y representación mediante el lenguaje natural. Gijón: Trea, 2004.

NEVES, D. A. B. *Aspectos metacognitivos da leitura do indexador*. 131 f. Tese (Doutorado em Ciência da Informação) – Universidade Federal de Minas Gerais, Belo Horizonte, 2004. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/EARM-73FMVG/doutorado___dulce_am_lia_de_brito_neves.pdf?sequence=1>. Acesso em 25 nov. 2015.

NEVES, D. A. B. Ciência da informação e cognição humana: uma abordagem do processamento da informação. *Ciência da Informação*, Brasília, v. 35, n. 1, p. 39-44, jan./abr. 2006.

NEVES, D. A. B. *Metacognição, informação e conhecimento*: pensando em como pensar. Recife: Nectar, 2011.

NEVES, D. A. B.; DIAS, E. W.; PINHEIRO, A. M. V. Uso de estratégias metacognitivas na leitura do indexador. *Ciência da Informação*, Brasília, v. 35, n. 3, p. 141-152, set./dez. 2006.

PINTO MOLINA, M. *Análisis documental*: fundamentos y procedimientos. 2. ed. rev. aum. Madrid: Eudema, 1993.

PINTO, M.; GÁLVEZ, C. *Análisis documental de contenido*: procesamiento de información. Madrid: Síntesis, 1999.

SAUSSURE, F. *Curso de linguística geral*. São Paulo: Cultrix, 1999.

SILVA, M. R.; FUJITA, M. S. L. A prática de indexação: análise da evolução de tendências metodológicas. *Transinformação*, Campinas, v. 6, n. 2, p. 133-161, maio/ago. 2004. Disponível em: <http://www.scielo.br/pdf/tinf/v16n2/03.pdf>. Acesso em: 05 out. 2016.

SOUZA, B. P.; FUJITA, M. S. L. Análise de assunto no processo de indexação: um percurso entre teoria e norma. *Informação & Sociedade: Estudos*, João Pessoa, v. 24, n. 1, p. 19-34, jan./abr. 2014. Disponível em: <file:///C:/Users/Usuario/Downloads/16281-36927-1-PB.pdf>. Acesso em: 25 out. 2015.

SOUZA, G. P. C. B. *A literatura infanto-juvenil brasileira vai muito bem, obrigada!* São Paulo: Difusão Cultural do Livro, 2006.

TUFANO, D.; NÓBREGA, M. J. *Esopo*: fábulas completas: projeto de leitura. São Paulo: Moderna, S.d. Disponível em: <http://www.moderna.com.br/lumis/portal/file/fileDownload.jsp?fileId=8A8A8A82375B303C01375C2796A11B73>. Acesso em: 10 dez. 2016.

VAN DIJK, T. A. *Cognição, discurso e interação*. São Paulo: Contexto, 2004. (Coleção Caminhos da Linguística).

VAN DIJK, T. A. El estudio del discurso. In: VAN DIJK, T. A. (Comp.). *El discurso como estructura y proceso*: estudios sobre el discurso I: uma introducción multidisciplinaria. Barcelona: Gedisa, 2000. p. 21-65. (Serie Cla-De-Ma Linguística/Análisis del Discurso).

VAN DIJK, T. A. *Estructuras y funciones del discurso*: una introducción interdisciplinaria a la lingüística del texto y a los estudios del discurso. 10. ed. Tradução Myra Gann e Martí Mur. Madrid: Siglo Veintiuno, 1996. Disponível em: <http://www.discursos.org/oldbooks/Teun%20A%20van%20Dijk%20-%20Las%20Estructuras%20y%20Funciones%20del%20Discurso.pdf>. Acesso em: 28 ago. 2015.

Os padrões de Hearst como recursos auxiliares semiautomáticos para a eficácia na leitura documentária

Walter Moreira
José Carlos Francisco dos Santos
Érica Fernanda Vitorini

1 Introdução

Como ponto de partida para a discussão que se faz neste capítulo, considera-se como imperativo a aplicação de metodologias de tratamento temático no amplo e variado conjunto da produção científica. A confluência das variáveis tempo destinado à análise dos documentos em sistemas de informação documentária e quantidade/qualidade destes documentos requer que o processo de tratamento temático seja realizado de modo eficaz, com o aproveitamento máximo dos recursos aplicados.

O tratamento temático da informação e, mais especificamente, a análise documentária é, por sua própria natureza intelectual, dispendioso. Desse modo, é fundamental que a análise documentária seja realizada com a observação atenta de princípios metodológicos e com o recurso de ferramentas que possam torná-la mais eficaz e racional. Entretanto, como adverte Kobashi (1994, p. 22), a análise documentária

> não se reduz a um conjunto de regras perenes, utilizáveis em todas as circunstâncias. Ela é, antes de tudo, uma disciplina de natureza metodológica que, para avançarteórica e praticamente, deve criticar continuamente seus pressupostos, procedimentos e instrumentos; deve, ao mesmo tempo, com base na reflexão permanente, elaborar novas hipóteses de trabalho que contribuam para aperfeiçoar os processos que lhe dizem respeito.

A análise documentária diz respeito ao exame do documento visando a sua representação para fins de recuperação da informação. Desse modo, a análise documentária configura-se como metodologia por meio da qual o conteúdo de um determinado texto é interpretado, condensado e lhe são atribuídos descritores, tendo em vista sua representação para fins de indexação.

A representação é um produto diferente do texto original, mas que guarda com ele relações de equivalência e de contiguidade. A construção da representação documentária é marcada por um processo de redução crescente, assim caminha-se

> do texto para o resumo, do resumo para o enunciado, do enunciado para a unidade de tradução via códigodocumentário. A atividade de AD caracteriza-se, portanto, como uma sucessão de processos de transformação do texto original, observando-se, a cada etapa, graus crescentes de generalização" (LARA, 1993, p. 41).

Nos modelos mais tradicionais dos sistemas documentários, esta representação documentária ocorre na forma de classificação, de indexação ou de resumo e tais representações visam a qualificar o processo de recuperação da informação pelo usuário.

Conforme Lara (2009, p. 28), o termo "análise documentária" foi criado por Jean-Claude Gardin para "designar as operaçõessemânticas que transformam um texto original em uma ou várias palavras-chave, ou ainda, paráfrases, visando facilitar a representação [...] de 'conteúdos' e a recuperação da informação".

No Brasil, as pesquisas desenvolvidas pelo Grupo Temma, que passa a existir formalmente com a publicação do livro "Análise documen-

tária: a análise da síntese", em 1987 (Lara, 2009), formam grande parte do estofo teórico e metodológico das pesquisas sobre análise documentária.

Para que se efetue a análise documentária, requer-se a realização de leitura documentária, uma modalidade de leitura profissional que consiste na análise do conteúdo do documento visando a identificação e a distinção das informações essenciais e acessórias, sendo que tal distinção ocorre conforme os interesses relativos ao contexto em que se produz a análise. Utilizando-se de técnicas de condensação documentária adequadas à modalidade de representação que se deseja construir, as informações essenciais são reelaboradas como representações documentárias e passam a compor, nessa condição, sistemas de informação documentária.

A leitura documentária, como qualquer atividade de leitura, é um processo de produção de sentido que tem como ponto de partida o texto. Neste caso há, contudo, uma variável que não está presente na leitura que se realiza para fins de lazer, por exemplo: o leitor profissional, o que realiza a leitura documentária, não é, normalmente, previsto como leitor pelo autor. Nas palavras de Lara (2009, p. 34): "frente ao documento, o leitor-documentalista não se caracteriza como um leitor-modelo, uma vez que não dispõe, necessariamente, de condições para estabelecer com o texto uma negociação". Além do mais, acrescenta-se o fator urgência (ou tempo reduzido), que obriga à realização de leitura mais rápida e seletiva do se poderia esperar a partir de um modelo convencional. Ainda citando Lara (2009, p. 34): "o leitor-documentalista realiza uma leitura que se enquadra num processo de produção industrial de textos (parafraseando Gardin), não podendo dedicar mais tempo à leitura do que aquele previsto na atividade de indexação de um grande volume de publicações".

Em busca de uma metodologia para a leitura documentária, Gil Leiva e Fujita (2012), estabelecem alguns procedimentos orientadores. Rubi, Fujita e Boccato (2012, p. 224) também caracterizam a leitura documentária e apresentam uma metodologia para a sua realização que envolve a análise mais cuidadosa dos seguintes elementos: a) introdução: com foco nos objetivos; b) "leitura das frases introdutórias de parágrafos e capítulos"; c) procedimentos metodológicos, contemplando "técnicas, instrumentos, procedimentos adotados na realização da pesquisa, bem como o e ambiência em que esta se passa"; d) conclusão em sua relação

com os objetivos propostos; e) "gráficos, tabelas, diferenciação tipográfica etc."; f) título, subtítulo, resumo e palavras-chave. Estes elementos devem ser verificados ao final da leitura, segundo as autoras, para que não a direcione de algum modo.

A leitura documentária é uma modalidade de leitura profissional, mais especificamente é a leitura realizada pelo indexador em contexto profissional. Trata-se de uma modalidade de leitura que visa a um fim específico que é construir a representação documentária. Não soa correto, portanto, do ponto de vista da organização lógica dos conceitos, tomar os termos "leitura documentária" e "leitura profissional" como sinônimos, em relação de equivalência. Desse modo, a expressão "leitura profissional", quando se refere à atividade realizada pelo indexador, deve aparecer adjetivada, como ocorre em Silveira e Moura (2007), com a expressão "bibliotecário-indexador" ou em Redígolo e Fujita (2015), que utilizam a expressão "leitura profissional do catalogador", cuja leitura profissional viabiliza a "consecução de seus objetivos de síntese e seleção de conceitos" (REDÍGOLO; FUJITA, 2015, p. 367)

Observado como leitor "comum", o indexador utiliza-se de recursos também "comuns" a qualquer outro leitor, acionando

> o processamento humano de informações, realizado com a memória de curto prazo (*input* visual), a memória de longo prazo (esquemas e conhecimento prévio) e as habilidades operatórias de pensamento (análise e síntese). [Observado como] leitor profissional é considerado a partir da perspectiva de seu contexto, atuação e formação profissional" (FUJITA; RUBI, 2006, p. 1).

Deste modo, são as técnicas empregadas na leitura que vão diferenciar o leitor "comum" do leitor profissional. De outro lado, é a tipologia de recursos empregados que irá diferenciar a leitura documentária no conjunto das leituras profissionais.

No trabalho em que sistematiza e apresenta uma abordagem teórico-metodológica para a construção de informações documentárias (especialmente o resumo), Kobashi (1994), quando discute a relação entre a prática documentária e o seu processamento automatizado, aponta que, naquele contexto, sem o conhecimento profundo das operações relativas

ao tratamento da informação documentária, os recursos da computação eram subutilizados. Citando-a textualmente, pode-se ler:

> Os fracassos [da incorporação da informática às tarefas da biblioteconomia/documentação] não se deveram unicamente à limitação das máquinas. As experiências revelaram que, muitas vezes, as dificuldades provinham de lacunas da própriaárea: especificamente, da ausência de conhecimentos suficientemente sistematizados sobre operações por vezes básicas e elementares do cotidiano profissional. Apoiada fortemente no conhecimento empírico, na intuição e no habito, a área, por um longo período, rejeitou a teoria, considerando-a supérflua.

O quadro não é mais o mesmo e houve evolução, naturalmente. Alguns autores, dentre eles Kobashi, no mesmo trabalho citado (1994) e Gil Leiva e Fujita (2012), forneceram importantes subsídios teóricos e metodológicos relativos aos processos inerentes à análise documentária. Sobre este aspecto, apontam-se apenas alguns outros trabalhos seminais: Smit (1989), Guimarães (1994), Lara (1999), Lucas (2000), Pinto Molina (2001), Fujita (2003), Montesi (2006) e Dias e Naves (2013).

A análise documentária, visa, como se disse anteriormente, a construção de representações documentárias que possam ser utilizadas tanto pelo indexador quanto pelo usuário pesquisador para fins de diálogos com o sistema. Neste caso, como se trata de uma linguagem construída para fins de comunicação entre usuários (leia-se indexador e pesquisador) e sistema, ou seja, uma linguagem documentária, é preciso que seja aplicado ao conjunto das representações documentárias metodologias de controle de vocabulário.

Para Lancaster (2002) o controle do vocabulário tem os objetivos de facilitar a representação dos assuntos tanto para os profissionais como para os usuários; agrupando os sinônimos e os quase sinônimos, diferenciando os homógrafos e relacionando os termos com significados próximos, sendo assim tem como fundamento diminuir essas diferenças e padronizar o seu uso facilitando, consequentemente, o acesso.

Deste modo, acredita-se que uma representação de qualidade pode proporcionar uma recuperação eficaz da informação. Problemas no processo de representação, como a ausência de um vocabulário controlado,

sua falta de atualização ou até mesmo um vocabulário que não represente a linguagem da comunidade usuária, geram falhas nessa comunicação e concorrem para a insatisfação na recuperação da informação. Além de comprometer a qualidade das buscas e do processo de recuperação da informação como um todo, a incompatibilidade entre a linguagem documentária e a linguagem do usuário, adverte Boccato (2009, p 21), compromete "a atuação do bibliotecário na representação dos conteúdos documentários no processo de indexação, o usuário na realização das buscas satisfatórias desses conteúdos no processo de recuperação da informação e, consequentemente, a credibilidade dos sistemas".

Embora se trate de um tema que não será explorado neste capítulo, é preciso reforçar o argumento de que não será possível estabelecer um diálogo eficaz entre usuários, linguagens documentárias, sistemas de informação documentária e indexadores (bibliotecários) sem a adoção de políticas de indexação coerentes e consistentes. A este respeito, pode-se consultar o já clássico artigo de Carneiro (1985) e os trabalhos de Rubi (2004; 2008), Gil Leiva e Fujita (2012) e Fujita (2016).

A política de indexação funciona como eixo orientador para o profissional indexador no momento da análise documentária, pois por meio dela é possível estabelecer o perfil da instituição e da comunidade usuária, entre outros critérios, e assim definir os parâmetros orientadores para a leitura e seleção dos descritores mais adequados à representação documentária.

A construção eficaz da representação como resultado da análise documentária requer, portanto, como condição *sine qua non*, um leitor orientado por uma política de indexação que: a) seja consciente a respeito da complexidade do ato de ler e que esteja munido de estratégias metacognitivas de leitura (um leitor profissional proficiente); b) conheça o perfil de interesse da comunidade usuária a que atende e c) conheça os diversos modelos de organização textual e saiba aplicar recursos adequados à identificação e seleção de informação essencial observando-se as diferentes tipologias de estrutura textual.

2 Organização da informação textual

O ponto de partida para a realização da análise documentária, conforme a compreensão do Grupo Temma relatada por Lara (2009, p. 30), é o texto, e não o discurso. A opção justifica-se pelo fato de se encontrarem no texto "as estruturas informacionais, elementos que respondem pela coesão e progressão textual, bem como pelo seu fechamento e autonomia".

Desse modo, é no texto que se verifica a condição de registro, o que lhe dá características de permanência no tempo e de portabilidade no espaço. Essas mesmas condições são apontadas como requisito para que o que vai ser definido como informação por Smit e Barreto (2002).

A análise documentária é constituída por três atividades específicas: análise, síntese e representação (Kobashi, 1994). As duas primeiras fases dizem respeito à desestruturação do texto e a última à sua reestruturação.

Na etapa de análise propriamente dita é que ocorre a leitura documentária, momento em que o leitor profissional (indexador) realiza a leitura do texto e identifica o seu assunto. Neste momento, a concepção clara do que representa a ideia de texto, de suas tipologias e de como se configuram suas estruturas são recursos essenciais ao pleno êxito da operação.

Um texto, na definição funcional apresentada por Dias e Naves (2013, p. 27), é tomado como meio, na condição de

> veículo que permite a comunicação de ideias entre o sujeito que cria e dissemina informação (emissor, no caso autor) e o sujeito que necessita e adquire informação(receptor, no caso o leitor). É o objeto que permite a transmissão das informações contidas em documentos, sendo também visto como uma coleção de símbolos, os quais são intencionalmente estruturados pelo emissor para mudar a estrutura de imagem do receptor.

Na fase de análise, o leitor profissional precisa lançar mão de estratégias de leitura que possibilitem a compreensão do texto de modo completo, preciso e rápido. Para isso, utiliza-se, essencialmente, "de representações mentais, que estão ligadas ao conhecimento linguístico, social e de representação" (REDÍGOLO; FUJITA, 2015, p. 357). Um texto é

composto por uma rede interna e externa de relações. É essa rede que, entre outros elementos, irá viabilizar a construção de sentido pelo leitor.

A compreensão do texto condiciona-se às melhores condições de interação entre o texto, o leitor e o contexto de ocorrência da atividade de leitura. Desse modo, é preciso reafirmar o papel ativo do leitor durante o processo, mormente no processo de leitura profissional.

Cintra (1989, p. 34) distingue dois tipos de estratégias relativas ao processo de leitura: estratégias cognitivas e estratégias metacognitivas. As primeiras "compreendem comportamentos automáticos e inconscientes", as últimas "supõem comportamentos desautomatizados, na medida em que o leitor tem consciência de como está lendo". O acionamento dos conhecimentos prévios como estratégia metacognitiva de leitura, por exemplo, dá ao leitor profissional uma perspectiva ampliada de análise, compreensão e representação do texto nas tarefas de leitura documentária.

O conhecimento prévio, também chamado algumas vezes de "conhecimento de mundo", refere-se a todo o conhecimento que o indivíduo possui, isto é, o conhecimento que o indivíduo armazenou na mente como resultado da capacidade inata de organização das suas experiências com o mundo (MEURER, 1985), diz respeito à "memória semântica" e à "memória episódica" do indivíduo, ou seja, a todo o conhecimento generalizado e particularizado armazenado na mente.

O acionamento do conhecimento prévio, bem entendido, não se configura como um mecanismo de acionamento "liga-e-desliga". Está presente em qualquer atividade de leitura. O modo como isso ocorre, de forma consciente ou não é que irá determinar a natureza das estratégias metacognitivas ou cognitivas de leitura. A considerar-se (hipoteticamente) um nível zero de conhecimento prévio, a interpretação de um objeto de natureza complexa como o texto escrito torna-se, mais do que difícil, literalmente desprovida de sentido e impraticável.

A eficácia da leitura depende também, como se disse anteriormente, da qualidade do texto, naturalmente. Neste caso, há variáveis intervenientes que independem do leitor (como o estilo de redação e os recursos argumentativos do autor) e há as que são perfeitamente apreensíveis, como

ocorre com a percepção consciente da estrutura textual e dos recursos linguísticos utilizados.

O sentido geral que se procura alcançar com o texto é baseado na sua estrutura, a qual compreende suas micro, macro e superestrutura. Dias e Naves (2013, p. 29) apresentam as seguintes definições para estes termos: microestrutura: "estrutura superficial, que corresponde a realidade física do texto e seus símbolos de significação, as palavras"; macroestrutura: "concebida como um tópico representativo hierárquico e coerente da unidade textual, envolvendo mínima estrutura da representação textual, sintático-semântica" e superestrutura. "estrutura retórico-esquemática, um tipo de esquema de produção convencional para o qual o texto é adaptado, podendo ser considerado como transição entre estruturas de superfície e de profundidade".

Nesta perspectiva, quanto mais conhecimento a respeitos dos gêneros e das estruturas textuais o leitor profissional tiver, maiores serão suas chances de identificar com clareza a ização das informações consideradas essenciais no texto, ou seja, maiores as possibilidades de análise e compreensão do texto visando a sua representação para fins documentários.

Considerando-se uma abordagem *top-down*, compreende-se o documento, objeto da leitura documentária, como uma complexa rede de relações conceituais. Em sentido inverso, em abordagem *bottom-up*, observa-se o conceito como unidade mínima de significação na comunicação de conhecimento que tem o texto como veículo. Trata-se de um "jogo" por meio do qual autor codifica as relações entre os conceitos que pretende tratar no texto (percurso onomasiológico, do enunciador) e estabelece comunicação com o leitor, cuja tarefa é decodificá-las (percurso semasiológico, do interpretante).

Não se pode ler profissionalmente e de modo proficiente, portanto, sem o recurso de estratégias metacognitivas de leitura. A leitura é um processo de interação entre texto, leitor e contexto e requer, para sua plena efetivação, a presença de um leitor ativo, consciente da sua função na construção de sentidos, na reconexão entre os conceitos.

3 Os padrões de Hearst

Marti A. Hearst, professora e pesquisadora na *School of Information, University of California*, Berkeley e que tem dentre seus temas de pesquisa interesses na linguística computacional, propõe em alguns textos (HEARST, 1992; 1998) métodos para a identificação automática e a utilização de padrões léxico-sintáticos como recursos para a expressão de relações léxico-semânticas.

As pesquisas de Hearst tomam como base grandes corpora de textos que estão disponíveis na internet para extração das informações lexicais, sintáticas e semânticas. O método de "extração de padrões léxico-sintáticos" (*Lexico-Syntactic Pattern Extraction* - LSPE), conforme apresentado em Hearst (1998, p. 1, tradução livre), "pretende ser útil como uma ajuda automatizada ou semiautomatizada para lexicógrafos e construtores de bases de conhecimento dependentes de domínio".

Os padrões de Hearst, respeitando-se a terminologia que a pesquisadora prefere utilizar, possibilitam a identificação de relações conceituais de hiperonímia e de hiponímia que ocorrem por meio das posições que tais conceitos assumem relativamente uns aos outros.

Uma análise etimológica breve dos termos deixa ver claramente seus significados. O prefixo "hiper" diz respeito a algo que está em posição superior (ou em excesso), como ocorre, por exemplo, em "hipertensão" (def.: "Med. Pressão excessiva exercida pelo sangue nas paredes dos vasos sanguíneos"); o prefixo "hipo", por sua vez, relaciona-se a algo que está em posição inferior (ou em escassez), como ocorre, por exemplo, em "hipotensão" (def.: "Med. Pressão do sangue nas paredes dos vasos sanguíneos inferior à normal; pressão baixa").

Assim, hiperonímia é a palavra que transmite o sentido do todo e hiponímia remete à ideia de parte, tipo ou item do todo. Um hiperônimo é um termo que está superordenado em relação a um hipônimo; este, por sua vez, está subordinado em relação ao hipônimo. Trata-se, como se vê, de uma relação assimétrica. O termo "árvores frutíferas" é hiperônimo de "macieiras", "abacateiros" e "mangueiras". Tomando-se a relação em sentido oposto, "macieiras", "abacateiros" e "mangueiras" são hipônimos de "árvores frutíferas". Trata-se também de um tipo de relação estrutural, ao

mesmo tempo em que "árvores frutíferas" é hiperônimo de "macieiras" pode ser também hipônimo de "árvores".

Para que se possa compreender com mais clareza a natureza dos padrões utilizados por Hearst, pode-se observar um dos exemplos que a autora apresenta (HEARST, 1998, p. 3, tradução livre): "Agar é uma substância preparada a partir de uma mistura de algas vermelhas, tais como Gelidium, para uso laboratorial ou industrial".

A maioria dos leitores desta frase (considerando-se o domínio específico em que se inserem este texto que ora se desenvolve e seus leitores potenciais), desconhece o significado do termo "gelidium". O modo como estão estruturadas as informações na frase, entretanto, permite inferir que "gelidium"' é uma espécie de "algas vermelhas". Neste caso, não ocorre uma definição clássica, não se está deliberadamente definindo o termo. O que possibilita a construção semântica é a presença de um padrão léxico--sintático, nomeadamente o padrão "tais como" que conecta os conceitos "algas vermelhas" e "gelidium".

Observe-se outro exemplo com a aplicação do mesmo padrão: "A arquivística é tratada como a disciplina que agrupa todos os princípios, normas e técnicas que regem as *funções de gestão dos arquivos*, **tais como** a *criação*, a *avaliação*, a *aquisição*, a *classificação*, a *descrição*, a *comunicação* e a *conservação*" (GARCIA; SCHUCH JÚNIOR, 2002, p. 46, grifos acrescentados). Neste caso, a organização léxico-sintática das informações possibilita a identificação de relações hierárquicas entre os conceitos envolvidos que estão mediados pelo padrão "tais como". Estas relações, utilizando-se o modelo de visualização normalmente adotado nos tesauros, pode ser apresentada da seguinte forma:

-----

funções de gestão de arquivos

TE criação

avaliação

aquisição

classificação

descrição

comunicação

conservação

Os padrões de Hearst, acredita-se, possuem potencial de aplicação na organização da leitura documentária. Tal potencial pode ser detectado nas características que Hearst (1998, p. 4, tradução livre) aponta relativamente ao conjunto de padrões léxico-sintáticos que indicam relações hierárquicas:

a) "eles ocorrem frequentemente em muitos gêneros textuais;
b) "eles (quase) sempre indicam a relação de interesse;
c) "eles podem ser reconhecidos com pouco ou nenhum conhecimento pré-codificado".

Os padrões identificados por Hearst (1998) em seu estudo relativo à língua inglesa, são descritos no Quadro 1.

Quadro 1 – *Os padrões de Hearst*

Padrão	Exemplo
NP_0 such as NP_1\{,NP_2 ... , (and | or) NP_i\} Obs.: NP = noun phrase ⇔ sintagma nominal	*Agar is a substance prepared from a mixture of red algae, such as Gelidium, for laboratory or industrial use.* *-- agar* *---- gelidum*
such NP_0 as \{NP_1,\}* \{(and | or)\} NP_2	*... works by such authors as Herrick, Goldsmith, and Shakespeare.* -- author ---- Herrick ---- Goldsmith ---- Shakespeare
NP_1 \{, NP_1\}* \{,\} or other NP_0	*Bruises, ..., broken bones or other injuries ...* *-- injury* *---- broken bones* *---- bruises*

NP_1 {, NP_2}* {,} and other NP_0	*... temples, treasuries, and other important civic buildings.* *-- civic buildings* *---- temples* *---- treasuries*
NP_0 {,} including { NP_1 ,}* {or | and} NP_2	*All common-law countries, including Canada and England ...* *-- comom-law countries* *---- Canada* *---- England*
NP_0 {,} especially { NP_1 ,}* {or | and} NP_2	*... most European countries, especially France, England, and Spain.* *-- european countries* *---- France* *---- England* *---- Spain*

Fonte: *Elaborado pelos autores com base em Hearst (1992; 1998)*

Os padrões de Hearst foram identificados para aquisição automática de relações lexicais hiponímicas em língua inglesa. Considerando-se, contudo, seu potencial de aplicação também em outros idiomas, alguns estudos têm sido desenvolvidos para a identificação de padrões léxico-sintáticos que evidenciam relações léxico-semânticas também em língua portuguesa.

Machado e Lima (2015) descrevem em sua pesquisa alguns trabalhos correlatos, tanto com aplicação em língua estrangeira, principalmente o inglês, quanto em língua portuguesa. Os autores fazem referências enfáticas aos trabalhos de Hearst (1992, 1998) e apontam de modo mais genérico os autores que utilizaram os padrões de Hearst como base para aplicações em línguas estrangeiras, entre eles estão Maedche e Staab (2002), Cederberg e Widdows (2003) e Degeratu e Hatzivassiloglou (2004). Machado e Lima (2015) também indicam trabalhos em que aparecem traduções dos padrões de Hearst visando a sua aplicação em língua portuguesa, tais como Freitas (2007), Freitas e Quental (2007), Baségio (2007), Taba (2013), Taba e Caseli (2014) e Machado (2015).

Não foram identificados trabalhos relativos aos padrões de Hearst com aplicações no domínio da ciência da informação. Os trabalhos citados

anteriormente que discutem a aplicação dos padrões de Hearst em língua portuguesa são academicamente relacionados às áreas da linguística ou da ciência da computação. A pesquisa de Freitas (2007), *e.g.*, está vinculada ao Departamento de Letras da Pontifícia Universidade Católica do Rio de Janeiro; Freitas e Quental (2007) publicaram seu trabalho no *Workshop* de Tecnologias da Informação e da Linguagem Humana; Taba e Caseli (2014) publicaram sua pesquisa no *International Conference on Language Resources and Evaluation*e o trabalho de Machado e Lima (2015) foi publicado na Revista Estudos da Linguagem. Os trabalhos mais próximos da ciência da computação são: Taba (2013), que desenvolveu sua pesquisa no Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, Baségio (2007) e Machado (2015), cujas pesquisas vinculam-se ao Programa de Pós-graduação em Ciência da Computação da Pontifícia Universidade Católica - RS.

Freitas (2007) optou por utilizar três padrões de Hearst e identificou adicionalmente três outros padrões com observações do *corpus* utilizado em sua pesquisa, como pode ser observado na Quadro 2.

Baségio (2007) trabalhou com os padrões de Hearst (1992) e fez uma adaptação na expressão "sintagma nominal" (*noun phrase* – NP) para substantivo (SU). Isto foi necessário, justifica o autor, porque a composição do corpus não permitia dispor da informação que era necessária, deste modo simplificou-se o sintagma nominal. O autor supracitado utiliza outros padrões além dos que foram propostos por Hearst (1992) e outros autores citados, porém não compõem o interesse deste estudo por se referirem de modo específico à língua francesa.

Taba (2013) trabalhou com dois dos padrões de Hearst, a partir das adequações realizadas por Freitas (2007) e Freitas e Quental (2007), como pode ser observado no Quadro 2. Taba (2013) utilizou de modo manual em seu estudo dois padrões criados por Freitas (2007) e Freitas e Quental (2007) e identificou, para os fins de sua análise e de seu *corpus*, a necessidade de criar mais três padrões. Estes padrões estão relacionados com a extração de hiponímia, considerada pelos autores como resultante da relação semântica "is-a".

Machado (2015) refere-se aos padrões como “regras”, e, como pode ser observado no Quadro 2, fez uso de cinco padrões de Hearst, observando as adaptações de Freitas (2007), Baségio (2007) e Taba (2013). A partir destes estudos Machado (2015) desenvolveu quatro “regras”.

Quadro 2 – *Os padrões de Hearst e suas adaptações*

<table>
<tr><th>Padrão Hearst (1992, 1998)</th><th>Freitas (2007) Freitas e Quental (2007)</th><th>Baségio (2007)</th><th>Taba (2013) Taba e Caseli (2014)</th><th>Machado (2015) Machado e Lima (2015)</th></tr>
<tr><td>NP$_0$ such as NP$_1${,NP$_2$ ... , (and | or) NP$_i$}</td><td>SN HHiper (tais como | como_PDEN) SN1 { , SN2 … , } (e | ou) SNi

SN Hiper, (tais como | como_ PDEN) SN1 { , SN2 ... ,}(e | ou) SNi</td><td>SUB como {(SUB,)*(ou|e)} SUB

SUB tal(is) como {(SUB,)*(ou|e)} SUB</td><td>SN_Hiper (tais como | como) SN {, SN} *(e|ou) SN</td><td>SN(,)? como (SN ,)*(SN (e|ou))*SN</td></tr>
<tr><td>such NP$_0$ as {NP$_1$,}* {(and | or)} NP$_2$</td><td>não utilizado</td><td>tal(is) SUB como {(SUB,)*(ou|e)} SUB</td><td>não utilizado</td><td>SN(,)? ta(is|l) como (SN ,)*(SN (e|ou))*SN</td></tr>
<tr><td>NP$_1$ {, NP$_1$}* {,} or other NP$_0$</td><td rowspan="2">SN HHipo { ,SN Hipoi } * { , } e|ou outros SN Hiper</td><td>SUB {, SUB}* {,} ou outro(s) SUB</td><td rowspan="2">SN {, SN}* , ? (e|ou) outros SN_Hiper</td><td rowspan="2">SN (ou|e|,))*<outr(a|o)(s)? sn></td></tr>
<tr><td>NP$_1$ {, NP$_2$}* {,} and other NP$_0$</td><td>SUB {, SUB}* {,} e outro(s) SUB</td></tr>
<tr><td>NP$_0$ {,} including { NP$_1$,}* {or | and} NP$_2$</td><td>não utilizado</td><td>SUB {,} incluindo {SUB,}*{ou|e} SUB</td><td>não utilizado</td><td>SN(,)? incluindo (SN ,)*(SN (e|ou))*SN</td></tr>
</table>

<table>
<tr>
<td>NP_0 {,}
especially { NP_1 ,}* {or | and}
NP_2</td>
<td>não utilizado</td>
<td>SUB {,}
especialmente
{SUB,}*{ou|e}
SUB

SUB {,}
principalmente
{SUB,}*{ou|e}
SUB

SUB {,}
particularmente
{SUB,}*{ou|e}
SUB

SUB {,} em
especial {
SUB,}*{ou|e}
SUB

SUB {,} em
particular {
SUB,}*{ou|e}
SUB

SUB {,} de
maneira especial
{ SUB,}*{ou|e}
SUB

SUB {,}
sobretudo {
SUB,}*{ou|e}
SUB</td>
<td>não utilizado</td>
<td>SN(,)?
especialmente
(SN ,)*(SN
(e|ou))*SN</td>
</tr>
<tr>
<td>-</td>
<td>tipos de SN
Hiper: SN1{
, SN2... ,} (e |
ou) SNi</td>
<td>-</td>
<td>tipos de SN_
Hiper: SN {,
SN}* (e|ou) SN</td>
<td><... tipo(s)?
de sn> : (SN
,)*(SN (e|ou)
)*SN</td>
</tr>
<tr>
<td>-</td>
<td>SN HHiper
chamado/s/a/as(
de) SN Hipo</td>
<td>-</td>
<td>SN_Hiper
chamad(o|a|os|as)
de? SN</td>
<td>SN(,| é| são|
foram)? chamad
(o|a|os|as)(de)?
(SN ,)* (SN
(e|ou))*SN</td>
</tr>
</table>

-	*SN Hiper *conhecido/s/a/as como* SN Hipo.*	-	-	SN((,)? também)? (,|é|são|foram)? conhecid (o|a|os|as) como (SN ,)*SN (e|ou))*SN”
-	-	-	SN {,SN} * ,? (e|ou) (qualquer| quaisquer) outro{s}? SN_ Hiper	(SN (ou|e|,))* < (qualquer| quaisquer) outr(a|o)(s)? sn>
-	-	-	SN é (o|a|um|uma) SN_Hiper	SN é < (o|a) sn> SN é < (um|uma) sn>
-	-	-	SN são SN_ Hiper	SN são SN

Fonte: *Elaborado pelos autores*

Cabe salientar que as pesquisas desenvolvidas levaram em consideração, como era de se esperar, os trabalhos previamente publicados. Desse modo, Machado (2015) utilizou o artigo da Taba e Caseli (2014) e Taba (2013) utilizou o artigo de Freitas e Quental (2007). Tendo-se isto em mente, neste estudo que se apresenta foram observados mais proximamente os trabalhos dos autores tomados como “seminais”: Freitas (2007), Baségio (2007), Taba (2013) e Machado (2015). Foram estes trabalhos que serviram de base para o desenvolvimento da fase experimental desta pesquisa e que subsidiaram as adaptações necessárias à análise.

A literatura atesta a potencialidade dos padrões de Hearst para a identificação de relações hierárquicas no texto. Esta literatura forma, contudo, um grande mosaico a respeito das questões particulares a cada contexto de análise. Os trabalhos realizados por Freitas (2007), Baségio (2007), Taba (2013) e Machado (2015), foram desenvolvidos a partir de *corpus* já analisados e etiquetados sintaticamente. Baségio (2007), Taba (2013) e Machado (2015) desenvolveram em suas pesquisas um protótipo de *software*. Em Taba (2013) este software é denominado “Anotador de Relações Semânticas (ARS)”.

Para fins de verificação de aplicabilidade dos padrões de Hearst no contexto da leitura documentária, observando-se de modo específico a identificação de termos e suas relações hierárquicas de modo semiautomático, foi realizado um experimento cujos procedimentos estão descritos na seção subsequente.

4 Procedimentos metodológicos

O experimento realizado neste estudo foi desenvolvido a partir de um *corpus* selecionado da Revista Latino-Americana de Enfermagem (RLAE) da Escola de Enfermagem de Ribeirão Preto - Universidade de São Paulo. Utilizou-se o seu volume mais recente (v. 25), contendo 26 artigos e um editorial. A escolha do referido periódico deu-se de modo aleatório, mas considerou-se sua qualificação no estrato "Qualis/Capes A1" da área de saúde, mais especificamente em enfermagem. Conforme informações disponíveis na página da RLAE, trata-se de um periódico bimestral que circula desde janeiro de 1993.

Para a realização da análise, foram executados os procedimentos de *download* dos textos em formato PDF e em língua portuguesa (a revista disponibiliza também versões em língua espanhola e inglesa dos artigos). A ferramenta escolhida para análise do corpus foi o *softwareWordSmith Tools* versão 4.0, com a licença *free*. Executou-se o procedimento de conversão para arquivos no formato TXT, legível pelo *WordSmith*, e realizou-se o tratamento manual dos textos para evitar a separação de parágrafos, já que acontece, muitas vezes, que na conversão de arquivos alguns parágrafos fiquem separados por quebras de páginas, figuras, tabelas, entre outros.

O *WordSmith Tools* é um conjunto integrado de funções para análise do comportamento das palavras nos textos; as principais funções são: *WordList*, *Concord* e *KeyWords*. O *WordList* possibilita visualizar uma lista de palavras ou um agrupamento. A função *Concord*, permite a ização de termos em textos e apresenta o contexto que aquele termo aparece, ou seja,possibilita listar as ocorrências e consultá-las na íntegra. O *KeyWords*possibilita a ização de um conjunto de palavras-chave nos textos. Além destas funções principais é possível obter outras informações estatísticas dos textos, como padrões, frequências, entre outras.

O *WordSmith Tools* foi desenvolvido por Mike Scott na *Oxford University Press* para o desenvolvimento de seu próprio trabalho lexicográfico e atende a diversos idiomas, inclusive o português.

Nesta pesquisa, trabalhou-se com duas modalidades de análise: a) semiautomática, a partir do uso da ferramenta *WordSmith* e b) manual. Na análise semiautomática foi utilizado o *corpus* em sua totalidade. Esta opção foi possível por conta do retorno rápido e preciso que a ferramenta tecnológica proporciona. Já na análise manual, foram utilizados os três primeiros artigos da revista.

A opção pela aplicação dos padrões de Hearst (1992, 1998) neste estudo foi fundamentada, como já se disse anteriormente, nos resultados dos trabalhos desenvolvidos por Freitas (2007), Baségio (2007), Taba (2013) e Machado (2015). Assim, foram utilizados os padrões descritos no Quadro 3, já com as devidas adaptações para efeitos de simplificação quanto ao seu entendimento e aplicação, mas mantendo-se, de qualquer modo, sua relação com os padrões de Hearst (1992, 1998). Como ocorre em quase toda simplificação, há, neste caso, alguma redução em relação ao algoritmo apresentado nos padrões originais.

Quadro 3 – *Os padrões léxico-sintáticos aplicados na pesquisa*

PADRÕES DE HEARST (1992, 1998)	PADRÕES APLICADOS
NP_0 such as NP_1{,NP_2 ... , (and | or) NP_i}	(1) SN (tais como | como) SN { , SN … , } (e | ou) SN
such NP_0 as {NP_1,}* {(and | or)} NP_2	(2) tal(is) SN como {(SN,)*(ou|e)} SN
NP_1 {, NP_1}* {,} or other NP_0	(3) SN {, SN}* {,} ou outro(s) SN
NP_1 {, NP_2}* {,} and other NP_0	(4) SN {, SN}* {,} e outro(s) SN
NP_0 {,} including { NP_1 ,}* {or | and} NP_2	(5) SN {,} incluindo {SN,}*{ou|e} SN
NP_0 {,} especially { NP_1 ,}* {or | and} NP_2	(6a) SN {,} especialmente {SN,}*{ou|e} SN (6b) SN {,} principalmente {SN,}*{ou|e} SN (6c) SN {,} particularmente {SN,}*{ou|e} SN (6d) SN {,} em especial { SN,}*{ou|e} SN (6e) SN {,} em particular { SN,}*{ou|e} SN (6f) SN {,} de maneira especial { SN,}*{ou|e} SN (6g) SN {,} sobretudo { SN,}*{ou|e} SN

Fonte: *Elaborado pelos autores*

Em relação aos procedimentos de aplicação, o algoritmo de Hearst (1992), é orientado pelos seguintes passos: 1) escolher a relação semântica; 2) obter uma lista de termos para validar as relações; 3) izar no corpus onde essas relações ocorrem sintaticamente próximas e registrar as ocorrências; 4) izar similaridades entre esses registros e a possível indicação da relação de interesse; 5) identificando-se positivamente um novo padrão, utilizá-lo para coletar mais instâncias da relação alvo e retornar à etapa 2.

Em relação ao passo 1, foram definidas as relações semânticas de hiperonímia e de hiponímia. Para a consecução do passo 2, utilizou-se a ferramenta *parserWordSmith*, anteriormente descrita, para izar no próprio *corpus* as relações que foram tomadas como relações positivas. Considerando-se que a lista de relações e seus termos foram extraídos do próprio *corpus* por meio da ização do termo-chave de cada padrão, os passos 3 e 4 foram executados juntamente com o passo 2.

Foram esgotadas as buscas por meio do termo-chave de cada um dos padrões e as ocorrências foram analisadas de maneira manual, homologando-se as hierarquias encontradas. A partir das identificações das relações, foram realizadas as análises visando ao atendimento do passo 5, isto é, avaliar positivamente as relações hierárquicas.

5 Identificação de estruturas hierárquicas com recurso dos padrões de Hearst

Os resultados estão estruturados tomando-se os padrões como categorias de análise. Apresentam-se, para cada padrão, excertos do *corpus*com a finalidade de ilustrar a sua aplicação. Apresentam-se adicionalmente, para os casos possíveis, a estrutura hierárquica conforme poderia ocorrer sua configuração em um tesauro.

Os exemplos apresentados variam em conformidade com o interesse da exposição dos resultados. Assim, tanto são apontados e discutidos casos em que os padrões permitiram a identificação de relações hierárquicas como casos em que isso não foi possível.

Na categorização das relações extraídas, verificou-se também algumas situações de atendimento parcial dos padrões, ou seja, casos em que

as ocorrências apresentavam algum conectivo não previsto entremeando os sintagmas nominais e os padrões.

Houve também casos em que não foi possível identificar a qualidade e a identidade lógica da relação hierárquica pela ausência de sentido geral no contexto do corpus analisado. Estes casos demandavam pesquisas como forma de homologar a relação hierárquica que não estão entre os objetivos deste trabalho.

Os resultados estatísticos gerais relativos às aplicações dos padrões de Hearst e seus índices de ocorrência no *corpus* podem ser verificados na Tabela 1, apresentada após as subseções que apresentam e discutem os padrões.

5.1 Padrão "tais como" / "como"

O padrão originalmente proposto por Hearst, adaptado e utilizado por Freitas (2007), Baségio (2007), Taba (2013) e Machado (2015), foi utilizado neste estudo. Este padrão é representado da seguinte forma:

- SN (tais como | como) SN { , SN … , } (e | ou) SN.

Aplicou-se parte do padrão "tais como" para recuperar os registros de ocorrência por meio da ferramenta *WordSmith*, obtendo-se como resultado 23 ocorrências no *corpus*. Posteriormente aplicou-se a consulta utilizando-se "como". Considerando-se que a relação lógica entre os componentes "tais como" e "como" neste padrão se dão pela presença de um "ou" exclusivo e observando-se que os resultados alcançados com o uso do "como" foram praticamente todos contemplados com o uso de "tais como", procedeu-se a exclusão dos registros recuperados em duplicidade, resultando, no total, 494 ocorrências.

Apresentam-se, na sequência, exemplos de aplicação do padrão "tais como | como", os quais podem ser visualizados por meio dos excertos 1 e 2 e de suas estruturas hierárquicas correspondentes. Estes dois exemplos, são considerados como positivos pelo sucesso na sua aplicação, isto é, pela extração de relação hierárquica pode meio da identificação dos

padrões léxico-sintáticos. O hiperônimo está izado no primeiro sintagma nominal antes da identificação do padrão (destacado em negrito) e os hipônimos logo depois.

Excerto 1 - "Mas alguns trabalhos rotineiros têm duplicações com os cuidados hospitalares, **tais como** infusões intravenosas, injeções e curativos".

-----

cuidados hospitalares

curativos

infusões intravenosas

injeções

Excerto 2 - "[...] Na atenção primária e em centros de atendimento ambulatorial, a atenção com a NP para pacientes com doenças crônicas, **como** doenças cardíacas, hipertensão e diabetes, resultou em melhores indicadores de controle das doenças [...]".

-----

doenças crônicas

diabetes

doenças cardíacas

hipertensão

O excerto 3 é uma amostra de caso em que o padrão não se aplica por não se verificar a ocorrência de sintagmas nominais, mas sim a presença de sintagma verbal antes do "como" previsto no padrão. Na verdade, a ocorrência "como", neste caso, tem aplicação completamente diversa do que prevê o padrão "como".

Excerto 3 - "[...] Seria mais apropriado empregar **como** denominador uma estimativa da exposição à gravidez de meninas de 10 a 14 anos de idade, no entanto estes dados não estavam disponíveis. [...]".

Vale observar que, em relação ao padrão "tais como", não foi verificada nenhuma ocorrência que impossibilitasse um enquadramento positivo, ainda que os aspectos semânticos das relações hierárquicas não sejam facilmente visíveis de modo automático. O excerto 4 ilustra este caso.

Excerto 4 - "[...] Na Parceria de Apoio, os profissionais de saúde apoiam os membros da comunidade conforme necessário de acordo com suas situações, **tais como** doenças ou condições de envelhecimento".

-----

situações de apoio da Parceria de Apoio

  condições de envelhecimento

  doenças

5.2 Padrão "tal(is)"

O padrão de Hearst denominado "tal(is)" não foi utilizado por Freitas (2007) e Taba (2013), mas aparece em Baségio (2007) e Machado (2015). Este padrão está representado como segue:

- tal(is) SN como {(SN,)*(ou|e)} SN.

O padrão "tal(is)" foi izado em apenas duas ocorrências. Em ambas não foi possível a extração positiva das relações hierárquicas, por suas inadequações em relação ao padrão. No excerto 5 é possível observar a satisfação parcial do que contempla o padrão.

Excerto 5 - "[...] Em tal contexto, o enfermeiro, **como** profissional de saúde, tem papel fundamental na elaboração e prática de intervenções que modifiquem essa realidade".

No excerto 6 há uma distância considerável entre "tal" e "como", o que inviabiliza o estabelecimento seguro de aplicação do padrão.

Excerto 6 - "[...] **Tal** categoria apreendeu conteúdo para cinco subcategorias que, no referido programa de computador, são traduzidas **como** codes".

5.3 Padrão "ou outro(s)"

Este padrão foi utilizado por Freitas (2007), que incorporou a este padrão o padrão "e outro(s)" (objeto da próxima subseção, de número 5.4). Do mesmo modo fizeram Taba (2013) e Machado (2015). Baségio (2007) tratou separadamente os dois padrões, do mesmo modo como ocorreu em relação a este trabalho de pesquisa. Este padrão apresenta-se do seguinte modo:

- SN {, SN}* {,} ou outro(s) SN.

Para o padrão "ou outro(s)", foram obtidas cinco ocorrências no texto, destas somente uma não compreende de modo positivo o padrão, demonstrada no excerto 7. Este exemplo reforça o argumento de que não se pretende, a partir do que é traçado nesta pesquisa, eleger abordagens completamente automáticas de leitura.

Excerto 7 - "[...] ou seleção aleatória, para os grupos, todas as participantes tiveram uma chance igual de serem incluídos em um **ou outro** grupo".

As demais ocorrências todas atendem ao padrão. No excerto 8, que ilustra um caso positivo, é possível observar que o hiperônimo está logo após o "ou outro(s)" e o hipônimo antes. Este exemplo ilustra também que não se pode dispensar, em qualquer caso, o recurso aos instrumentos de controle de vocabulário, como os tesauros. Neste caso, o próprio texto faz menção à substituição da expressão "líquidos corporais" por "fluidos corporais".

Excerto 8 - "[...] Eu limpo imediatamente com desinfetante (álcool) superfícies após derramamento de sangue **ou outros** líquidos corporais.

-----

fluidos corporais

    sangue

5.4 Padrão "e outro(s)"

Este padrão é representado como segue:

- SN {, SN}* {,} e outro(s) SN.

No total, foram izadas no corpus, 22 ocorrências relativas a este padrão. Destas, duas não atenderam, três atenderam parcialmente e 17 atenderam de modo positivo a aplicação do padrão em destaque.

Apresenta-se no excerto 9 um dos casos em que a aplicação do padrão ocorreu de modo positivo.

Excerto 9 - "[...] É impossível planejar a APS sem ter claros os papéis para os médicos, enfermeiras, parteiras **e outros** profissionais de saúde.

-----

profissionais da saúde

enfermeiros

médicos

parteiros

Pode-se observar que hiperônimo fica izado após a ocorrência do padrão "e outro(s)" e os hipônimos antes dele, de modo semelhante ao que ocorreu com o padrão "ou outro(s)".

O excerto 10 demonstra o não atendimento ao padrão. A expressão "e outro", neste caso, é recurso estilístico para diferenciar os dois tipos de agentes envolvidos no que se relata, sem que apresentem relações de hierarquia entre si.

Excerto 10 - "[...] Além delas, foram escolhidos dois profissionais – um que realizou as intervenções comportamentais **e outro** que realizou as intervenções educativas por telefone".

5.5 Padrão "incluindo"

Este padrão não foi utilizado por Freitas (2007) e Taba (2013). Baségio (2007) e Machado (2015), entretanto, aplicaram-no em suas pesquisas. Eis sua representação:

- SN {,} incluindo {SN,}*{ou|e} SN.

Foram identificadas 57 ocorrências deste padrão no *corpus*. Em trinta delas, houve total compatibilidade com o que se esperava da aplicação do padrão, isto é, permitiram a identificação de hiperônimos e hipônimos. Em dezenove ocorrências houve resposta parcial e em oito casos a expressão identificada guardava apenas relação de similaridade com o padrão, ou seja, não havia relações hierárquicas entre os termos que a expressão tomada como padrão mediava.

No excerto 11 é possível observar a relação hierárquica identificada com a aplicação do padrão "incluindo". Identificam-se o hiperônimo como seu antecessor e os hipônimos como sucessores.

Excerto 11 - "[...] Resultados: participaram do estudo 573 profissionais, **incluindo** técnicos e auxiliares de enfermagem 292 (51%), enfermeiros 105 (18,3%), médicos 59 (10,3%), e outros profissionais 117 (20,4%)".

-----

profissionais

  auxiliares de enfermagem

  enfermeiros

  médicos

técnicos de enfermagem

Cabe mais uma vez a advertência: os casos são apresentados como exemplos e procurou-se manter o máximo possível de fidelidade em re-

lação ao contexto de ocorrência para que fosse possível compreender a aplicação dos padrões de Hearst. O termo "profissionais" no excerto 11, por exemplo, mantido na estrutura hierárquica apresentada, é vago. Além do próprio contexto de ocorrência do padrão (uma revista de enfermagem) informar, há menção explícita no texto aos "profissionais da saúde". Independentemente disto, contudo, a utilização de um tesauro como recurso para a tradução no processo de indexação resolveria, pela sugestão de termos mais precisos, a questão.

O exemplo apresentado no excerto 12 é um dos casos em que a relação hierárquica não pôde ser identificada plenamente e de modo formal pela aplicação automática do padrão "incluindo". Um leitor profissional, ou mesmo um outro tipo de leitor proficiente, percebe facilmente que as expressões que aparecem depois da palavra "incluindo" ("utilização de EPI", "uso de sistemas de exaustão " e "ventilação eficaz nas salas operatórias") são relativas às "medidas preventivas nos CC para a minimização dos riscos químicos devido à exposição à inalação da fumaça cirúrgica", funcionando com seus hipônimos. Ocorre que a expressão é muito longa para funcionar de modo adequado como descritor e nem mesmo a utilização pura e simples de um vocabulário controlado pode resolver prontamente a questão.

Excerto 12 - "[...] Portanto, tem-se como prioridade a adoção de medidas preventivas nos CC para a minimização dos riscos químicos devido à exposição à inalação da fumaça cirúrgica, incluindo a utilização de EPI e o uso de sistemas de exaustão e de ventilação eficaz nas salas operatórias".

O que se busca identificar nesta pesquisa, rememora-se, é a identificação de relações semânticas hierárquicas de modo semiautomático com o recurso de padrões léxico-sintáticos. Neste caso específico em análise, a resposta não é completamente satisfatória, justamente porque a tarefa maior de identificação dos sentidos atribuídos aos termos fica por conta do leitor, o que torna a utilização do padrão com recurso de leitura no mínimo discutível.

5.6 Padrão "especialmente" e semelhantes

Este padrão foi utilizado por Baségio (2007) e Machado (2015). Freitas (2007) e Taba (2013) não o aplicaram em suas pesquisas. Machado (2015) fez uso apenas do padrão "especialmente", Baségio (2007) ramificou este padrão para "principalmente", "particularmente", "em especial", "em particular", "de maneira especial" e "sobretudo". Para os efeitos desta investigação, optou-se pela aplicação e análise de todos estes padrões.

Como resultado geral da análise, verificou-se que os padrões "especialmente", "principalmente", "particularmente", "em especial", "em particular", "de maneira especial" e "sobretudo" são menos precisos, de modo geral, que os demais. No interior deste grupo, apenas o padrão "especialmente" destaca-se dos demais.

O índice de ocorrências positivas para este grupo de padrões foi de 29 acertos. Destes, o padrão "especialmente" responde por vinte acertos (em 39 ocorrências) e o padrão "principalmente" por quatro acertos (em 22 ocorrências). Considerando-se isto, optou-se por descrever apenas os resultados destes últimos padrões.

O padrão "especialmente" (e suas variações), pode ser representado do seguinte modo:

- SN {,} especialmente {SN,}*{ou|e} SN.

Como resultado da aplicação do padrão "especialmente", foram identificadas 39 ocorrências, sendo que vinte satisfazem totalmente ao que se espera do padrão, onze atendem parcialmente e oito não atendem, isto é, não possibilitam a extração de relações hierárquicas.

No excerto 13 é possível observar o hiperônimo, que antecede ao padrão "especialmente", e seus respectivos hipônimos.

Excerto 13 - "[...] São necessários mais serviços de promoção da saúde, prevenção e gestão para reduzir a carga de doença e a mortalidade associada a doenças crônicas, **especialmente** a saúde mental, câncer, doenças cardiovasculares e diabetes".

-----

doenças crônicas

câncer

diabetes

doenças cardiovasculares

saúde mental

Neste caso, embora a aplicação e o resultado obtido com o padrão estejam qualificados como positivos, percebe-se claramente um erro lógico causado pela ambiguidade da redação do texto original: "saúde mental" não é, evidentemente, hipônimo de "doenças crônicas", talvez a ausência desta o seja. Conforme já dito anteriormente, procurou-se ater de modo específico ao corpus justamente para que se pudesse observar o comportamento na aplicação dos padrões de Hearst. Além disso, é importante lembrar que a qualidade da relação lógica não pode ser depreendida de um ou outro excerto, de um ou outro *corpus*, mas sim de uma verificação sistemática deles.

O excerto 14, reproduzido na sequência, ilustra um caso em que não há resposta satisfatória com a aplicação do padrão. A presença de sintagmas verbais antecedendo o padrão "especialmente" dificulta, neste caso, a extração confiável das relações hierárquicas.

Excerto 14 - "[...] Para suprimir as restrições financeiras de cuidados com a saúde, é importante melhorar o estilo de vida para prevenir as DNT e promover saúde, reforçando **especialmente** o conhecimento e alfabetização em saúde para todas as idades.

Em relação ao padrão "principalmente", uma derivação ou variação do padrão "especialmente", foram identificadas 22 ocorrências no *corpus*: quatro positivas, dez parciais e oito em que a identificação do padrão não apresentou resposta satisfatória.

No excerto 15 reproduz-se um caso positivo em que se verifica uma relação meronímica.

Excerto 15 - "[...] Desse modo, torna-se necessário o desenvolvimento de pesquisas para que se possam obter melhores conhecimentos sobre o ambiente de trabalho, **principalmente** nos serviços de emergência e como o mesmo interfere na prática profissional.

-----

ambientes de trabalho

serviços de emergência

Na Tabela 1, que apresenta os resultados estatísticos gerais relativos às aplicações dos padrões de Hearst e seus índices de ocorrência no *corpus, foram*consideradas ocorrências todas as instâncias das expressões indicativas do padrão.

Tabela 1–*Distribuição geral dos padrões no corpus*

Nº	PADRÃO	OCORRÊNCIAS	RELAÇÕES	%	ACERTOS	%
1	SN (tais como | como) SN { , SN … , } (e | ou) SN	518	228	44,02	130	25,1
2	tal(is) SN como {(SN,)*(ou|e)} SN	2	1	50	-	-
3	SN {, SN}* {,} ou outro(s) SN	5	4	80	4	80
4	SN {, SN}* {,} e outro(s) SN	22	20	90,91	17	77,27
5	SN {,} incluindo {SN,}*{ou|e} SN	57	49	85,96	30	52,63
6a	SN {,} especialmente {SN,}*{ou|e} SN	39	31	79,49	20	51,28
6b	SN {,} principalmente {SN,}*{ou|e} SN	22	14	63,64	4	18,18

6c	SN {,} particularmente {SN,}*{ou|e} SN	2	2	100	1	50
6d	SN {,} em especial { SN,}*{ou|e} SN	1	1	100	1	100
6e	SN {,} em particular { SN,}*{ou|e} SN	6	2	33,33	-	-
6f	SN {,} de maneira especial { SN,}*{ou|e} SN	0	-	-	-	-
6g	SN {,} sobretudo { SN,}*{ou|e} SN	10	7	70	3	30
	TOTAL	**684**	**359**	**52,49**	**210**	**30,7**

Fonte: *Elaborada pelos autores*

A coluna "relações" refere-se às quantidades de relações positivas extraídas, incluindo-se as que atenderam parcialmente ao padrão. Considerou-se "acertos" todas as relações positivas, isto é, casos em que foi possível identificar relações semânticas a partir da identificação de relações lexicais.

Considerações finais

Pode-se considerar que há dois macro-processos orientados ao tratamento da informação encontrada em conjuntos textuais: a extração e a abstração. O primeiro relaciona-se aos padrões léxico-sintáticos, refere-se ao que está presente no texto, o segundo diz respeito ao que não está formalizado no texto. Os computadores, sabe-se, são mais eficientes na primeira tarefa, isto é, produzem resultados mais confiáveis quando aplicados em atividades de extração do que em atividades de abstração.

A atividade de leitura documentária para fins de indexação é, pelo conjunto de variáveis que envolve (já apresentadas anteriormente), extremamente complexa. Trata-se de uma atividade que demanda algum nível de extração, mas cuja natureza configura-se como essencialmente abstrata. Indexa-se não o que está no texto, ou pelo menos não necessariamente o que está no texto, mas sim o que não está nele. O que se quer identificar e representar com processo de leitura documentária são os conceitos; o acesso aos conceitos (o que está ausente), entretanto, é de

natureza representacional, intermediado pelos termos (o que está presente) nos textos científicos.

Os aspectos teóricos e os resultados apontados neste trabalho apontam para dois aspectos distintos e inter-relacionados referentes à leitura documentária: os quase insondáveis e complexos aspectos sociocognitivos que a envolvem, e formam o seu núcleo, e alguns aspectos formais que, a partir do conhecimento da estrutura textual do documento que se analisa e do recurso a algumas ferramentas adequadas, podem ser identificados a partir de determinados padrões léxico-sintáticos que correspondem a determinados "padrões" léxico-semânticos.

É preciso reafirmar que, conforme aponta este estudo e também os que foram apontados ao longo do trabalho, não existem, a rigor, padrões universais (nem se pretendia encontrá-los, aliás). Para cada cultura, para cada domínio, para cada corpus estabelece-se um padrão próprio que pode, como é próprio dos artefatos culturais, alterar-se. A noção de padrão, portanto, deve ser formada a partir de análises periódicas sistemáticas, originadas de *corpus* representativos dos domínios que se pretende representar.

Dada esta complexidade, nem sempre os padrões funcionam, ou melhor, nem sempre as expressões tomadas como padrão são empregadas da mesma forma. A língua, mesmo nas terminologias, não é apenas código. Por este motivo, não se insinua, em nenhum momento, o recurso aos padrões léxico-sintáticos como substituto ao leitor profissional, o que seria, nestas condições, ingenuidade, para dizer o mínimo. Compreendeu-se aplicação de tais recursos como relativos a uma forma de abordagem semiautomática, que visa a instrumentalizar o processo de leitura documentária.

Os usos dos padrões léxico-sintáticos indicados no estudo não estão livres das ambiguidades e das alterações de sentido que podem sofrer em função dos recursos estilísticos de quem os emprega. O padrão "tais como", por exemplo, foi identificado nos resultados como um dos que mais positivamente responderam à identificação de relações hierárquicas às quais entremeava. A expressão "tais como", entretanto pode ser perfeitamente empregada para indicar relações de outra natureza, como – para empregar um último exemplo – a que se pode observar na citação extraída

de Assumpção (2011, p. 95, grifos acrescentados): “Na literatura é possível encontrar algumas variaçõesterminológicas para denominar um arquivo de autoridade, tais como ‘catálogo de autoridades’, ‘lista de cabeçalhos autorizados’, ‘lista de autoridades’, ‘catálogo de identidade’, ‘catálogo de formas autorizadas’, entre outros”. O leitor profissional, atento e com algum conhecimento acerca do domínio, percebe que não são de natureza hierárquica as relações entre os termos envolvidos. Essas relações referem-se a relações de equivalência: “catálogos de autoridade”, “lista de cabeçalhos autorizados” etc. não são tipos-de ou parte-de “arquivos de autoridade”, são antes variações terminológicas (ou sinonímicas), como está, aliás, explicitamente indicado no texto.

Acredita-se, por fim, que o acréscimo contínuo de textos em formato eletrônico e o incremento do volume de hipertextos não apenas requerem como possibilitam, em função de seus formatos, estruturas e suportes, o tratamento semiautomático da identificação de informações textualmente e contextualmente importantes. Às habilidades tradicionalmente requeridas ao bibliotecário relativas às operações da análise documentária devem ser acrescentados conhecimentos oriundos das ciências cognitivas, da linguística textual e da análise do discurso, entre outras.

Este estudo limitou-se à análise da aplicação dos padrões de Hearst a um *corpus* relativamente pequeno, visando testar a viabilidade de aplicação dos padrões léxico-sintáticos para identificação de relações léxico-semânticas. O volume do *corpus* não lhe dá caráter de representatividade em relação á área da enfermagem.

Conclui-se que a aplicação de dos padrões léxico-sintáticos selecionados para o estudo é potencialmente útil como recurso semiautomático para a realização da leitura documentária. Em estudos futuros, pretende-se ampliar o *corpus*incluindo outros domínios e o conjunto de padrões para que possam incluir, além das relações de hiperonímia e hiponímia, outras relações lexicais, como a meronímia e a holonímia, as relações de equivalência e as sempre mais complexas relações associativas.

Referências

ASSUMPÇÃO, F. A. *Controle de autoridade*: definições, processos e componentes. Monografia (Bacharelado em Biblioteconomia) - Universidade Estadual Paulista, Marília, 2011.

BASÉGIO, T. L. *Uma abordagem semi-automática para identificação de estruturas ontológicas a partir de textos na língua portuguesa do Brasil*.Dissertação (Mestrado em Ciência da Computação) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2007.

BOCCATO, V. R. C. *Avaliação do uso de linguagem documentária em catálogos coletivos de bibliotecas universitárias*: um estudo sociocognitivo com protocolo verbal. 2009. 301 f. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista, Marília, 2009.

CARNEIRO, M. V. Diretrizes para uma política de indexação. *Revista da Escola de Biblioteconomia da UFMG*, v. 14, n. 2, p. 221-241, set. 1985.

CEDERBERG, S.; WIDDOWS, D. Using LSA and noun coordination information to improve the precision and recall of automatic hyponymy extraction. In: CONFERENCE ON COMPUTATIONAL NATURAL LANGUAGE LEARNING, 7, 2003, Edmonton. *Proceedings...* Edmonton: Association for Computational Linguistics, 2003. CONLL, v. 4 p. 111-118.

CINTRA, A. M. M. Estratégias de leitura em documentação. In: SMIT, J. W. (Coord.). *Análise documentária*: a análise da síntese. 2.ed. Brasília: IBICT, 1989. p. 63-87.

DEGERATU, M.; HATZIVASSILOGLOU, V. An automatic method for constructing domain-specific ontology resources. In: LANGUAGE RESOURCES AND EVALUATION CONFERENCE, 4, 2004, Lisboa. *Proceedings...* Lisboa, 2004. p. 2001-2004.

DIAS, E. W.; NAVES, M. M. L. *Análise de assunto*: teoria e prática. 2.ed.rev. Brasília: Briquet de Lemos, 2013.

FREITAS, M. C.; QUENTAL, V. S. D. B. Subsídios para a elaboração automática de taxonomias. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, 27.; Workshop em Tecnologia da Informação e da Linguagem Humana, 5., 2007, Rio de Janeiro. *Anais...* Rio de Janeiro: SBC, 2007.

FREITAS, M. C. *Elaboração automática de ontologias de domínio*: discussão e resultados.Tese (Doutorado em Letras). Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2007.

FUJITA, M. S. L. (Org.). *Política de indexação para bibliotecas*: elaboração, avaliação e implantação. Marília: Oficina Universitária, 2016.

FUJITA, M. S. L. *A leitura documentária do indexador*: aspectos cognitivos e linguísticos influentes na formação do leitor profissional. Tese (Livre-docência) - Universidade Estadual Paulista Júlio de Mesquita Filho, 2003.

FUJITA, M. S. L.; RUBI, M. P. Um modelo de leitura documentária para a indexação de artigos científicos: princípios de elaboração e uso para a formação de indexadores. *DataGramaZero*, v. 7, n. 3, jun. 2006.

GARCIA, O. M. C.; SCHUCH JÚNIOR, V. F. A aplicação da arquivística integrada, considerando os desdobramentos do processo a partir da classificação. *Informação & Informação*, v. 7, n. 1, p. 41-56, jan./jun. 2002.

GIL LEIVA, I.; FUJITA, M. S. L. (Eds.). *Política de indexação*. Marília: Oficina Universitária, 2012.

GUIMARÃES, J. A. C. *Análise documentária em jurisprudência*: elementos para uma metodologia de indexação de acórdãos trabalhistas brasileiros. Tese (Doutorado em Ciências da Comunicação) - Universidade de São Paulo, 1994.

HEARST, M. A. Automated discovery of WordNet relations. In: FELLBAUM, C. (Ed.). *WordNet*: an electronic lexical database and some of its applications. Cambridge: MIT Press, 1998.

HEARST, M. Automatic acquisition of hyponyms from large text corpora. In: CONFERENCE ON COMPUTATIONAL LINGUISTICS, 14., Nantes, 1992. *Proceedings…* Nantes, 1992.

KOBASHI, N. Y**.** *A elaboração de informações documentárias*: em busca de uma metodologia. Tese (Doutorado em Ciências da Comunicação) – Universidade de São Paulo, São Paulo, 1994.

LANCASTER, F. W. *El control del vocabulario en la recuperación de información*. 2 ed. València: Universitat de València, 2002.

LARA, M. L. G. *A representação documentária*: em jogo a significação. Dissertação (Mestrado em Ciências da Comunicação) – Universidade de São Paulo, São Paulo, 1993.

LARA, M. L. G. *Linguística documentária*: seleção de conceitos. Tese (Livre-docência) – Universidade de São Paulo, São Paulo, 2009.

LARA, M. L. G. *Representação e linguagens documentárias*: bases teórico-metodológicas. Tese (Doutorado em Ciências da Comunicação) – Universidade de São Paulo, São Paulo, 1999.

LUCAS, C. R. *Leitura e interpretação em Biblioteconomia*. Campinas: Unicamp, 2000.

MACHADO, P. N.; LIMA, V. L. S.. Extração de relações hiponímicas em um corpus de língua portuguesa. *Revista de Estudos da Linguagem*, v. 23, n. 3, p. 599-640, dez. 2015.

MACHADO, P. N. *Extração de relações hiponímicas em corpora de língua portuguesa*. Dissertação (Mestrado em Ciência da Computação) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2015.

MAEDCHE, A.; STAAB, S. *Ontology learning for the semantic web*. Massachusetts: Kluwer Academic, 2002.

MEURER, J. L. Schemata in reading comprehension. *Ilha do Desterro*, n. 13, p. 31-46, 1985.

MONTESI, M. *Método de evaluación y calidad de resúmenes documentales*. Gijón: Trea, 2006.

PINTO MOLINA, M. *El resumen documental*: principios y métodos. Madrid: Fundación Germán Sánchez Ruipérez, 2001.

REDIGOLO, F. M.; FUJITA, M. S. L. A leitura profissional do catalogador e seu papel como mediadora da informação. *Informação & Informação*, v. 20, n. 3, p. 356 - 376, set./dez. 2015.

RUBI, M. P. *A política de indexação na perspectiva do conhecimento organizacional*. Dissertação (Mestrado em Ciência da Informação) - Universidade Estadual Paulista, 2004.

RUBI, M. P. *Política de indexação para construção de catálogos coletivos em bibliotecas universitárias*. Tese (Doutorado em Ciência da Informação) - Universidade Estadual Paulista, 2008.

RUBI, M. P.; FUJITA, M. S. L.; BOCCATO, V. R. C. Elaboração do manual de política de indexação na formação continuada do catalogador. In: GIL LEIVA, I.; FUJITA, M. S. L. (Eds.). *Política de indexação*. Marília: Oficina Universitária, 2012. p. 217-227.

SILVEIRA, F. J. N.; MOURA, M. A. A estética da recepção e as práticas de leitura do bibliotecário-indexador. *Perspectivas em Ciência da Informação*, v. 12, n. 1, p. 125-135, jan.abr. 2007.

SMIT, J. W. (Coord.). *Análise documentária*: a análise da síntese. 2.ed. Brasília: IBICT, 1989.

SMIT, J.; BARRETO, A. A. Ciência da informação: base conceitual para a formação do profissional. In: VALENTIM, M. L. P. (Org.). *Formação do profissional da informação*. São Paulo: Polis, 2002. p. 9-23.

TABA, L. S.; CASELI, H. M. *Automatic semantic relation extraction from portuguese texts*, 2014. Disponivel em: <http://www.lrec-conf.org/proceedings/lrec2014/pdf/522_Paper.pdf>. Acesso em: 02 abr. 2017.

TABA, L. S. *Extração automática de relações semânticas a partir de textos escritos em português do Brasil*. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de São Carlos, São Carlos, 2013.

Información visualizada: leyendo gráficos, tablas e infografías

María del Carmen Agustín Lacruz
Juan-Francisco Torregrosa Carmona

1. El nuevo entorno de la información: mass media y comunicación científica.

Los seres humanos hemos sido forjados por la evolución para ser una especie visual. Una parte de nuestro cerebro está dedicada a procesar información que captamos a través de los ojos. No tiene sentido que no aprovechemos esa capacidad para comunicarnos con más eficacia, rigor y originalidad. Esto puede parecer obvio, pero no lo es. La imagen tiene mucha importancia en la comunicación, porque entendemos mejor lo que vemos que lo que nos cuentan. Durante siglos, la imagen como forma de comunicación y conocimiento fue relegada a un espacio secundario, menos noble, que el texto. Y sin embargo, parece muy lógica la frase que se atribuye a Goethe: "El órgano con el que he comprendido el mundo ha sido el ojo".

A modo de ejemplo de esa infravaloración de la ilustración frente al texto, también en el ámbito de la prensa hasta fechas recientes, baste recordar que no fue hasta el 27 de diciembre de 1983 cuando el prestigioso periódico francés *Le Monde* rompió su tradición de no publicar

reproducciones en color como ilustración de su primera página. Un consejo de redacción decidiría entonces acabar con esa costumbre publicando el cuadro *Mujer ante el sol,*en el que un gran astro rojo figura como detalle rotundo de la obra del artista catalán Joan Miró (1893-1983)[1]. Fue un homenaje al artista, uno de los grandes representantes del surrealismo pictórico, que había fallecido el día 25 de ese mes de diciembre.

Este ejemplo demuestra que, incluso cuando era posible técnicamente aprovechar el potencial cromático de la imagen, existían recelos al considerar que una ilustración en color no tenía cabida en una publicación diaria de prensa de calidad por su supuesto nivel inferior de seriedad respecto al blanco y negro de siempre. No es menos cierto tampoco que, hoy como ayer, cualquier reproducción en colores encarece de forma notable el producto editorial en el que se incluya.

Fig 1. *Mujer ante el sol.* Joan Miró.

Treinta y cinco años después, la situación es radicalmente distinta. Hoy es la imagen, fotográfica o de otra naturaleza, la que gana terreno al texto, tanto en la comunicación periodística convencional (impresa) como sobre todo en las nuevas publicaciones digitales que dominan el

[1] *El País*, edición impresa del 29 de diciembre de 1983. Disponible en http://elpais.com/diario/1983/12/29/cultura/441500402_850215.html [Consultado el 18 de abril de 2017].

panorama informativo de Internet en todo el mundo industrializado, con una presencia creciente del vídeo y de la imagen en movimiento como elemento primordial.

En ese mismo contexto, nos enfrentamos en estos momentos al enorme desafío que supone la sobreabundancia y la complejidad de los datos y de la información. Ante esta situación, los seres humanos hemos ideado numerosas técnicas y herramientas para representar de forma gráfica, bella y comprensible la información (ALCALDE, 2015, p.10-13).

De hecho, muchas veces ni siquiera leemos la información que recuperamos en Internet y ojeamos a través de la pantalla. Saltamos de un sitio a otro leyendo solo lo que nuestro cerebro, a través de nuestros ojos, considera interesante. Esta nueva forma de consumir información nos obliga a sintetizar, esquematizar y a disponer la información de la forma más atractiva posible y así provocar una mejor comprensión, asimilación y retención en la memoria.

Por ello, el aporte documental de gráficos e infografías puede ser muy valioso valioso cuando aparece sintetizado y presentado de la forma adecuada y eficaz a la que están acostumbrados los lectores en el ámbito de la red, especialmente para los denominados *nativos digitales*, pues las generaciones más jóvenes prácticamente no se plantean acceder a los soportes tradicionales, a los que no consideran vehículos idóneos para recuperar información.

Es una realidad que los diarios convencionales pierden lectores, mientras que los ganan estas mismas cabeceras de prensa y otras nuevas en el vigente y muy reñido –desde el punto de vista de la atención, la influencia social y la rentabilidad económica– ecosistema digital.

El llamado *periodismo de datos*, dentro de la investigación informativa, es un ámbito de particular interés para la visualización y la comunicación gráfica.

En el caso de la comunicación científica, Cairo (2008) encuentra en el pensamiento y la polifacética obra artística de Leonardo da Vinci su idea central sobre la ilustración al servicio de la ciencia y por extensión –entendemos– de la cultura y la investigación: "las imágenes que los científicos crean, usan y manipulan no son meros aderezos, simples apoyos

visuales de argumentos tejidos con palabras; en ciencia, las imágenes *son* pensamiento" (la cursiva es del autor).

2. Arquitectura, diseño y visualización de la información: conceptos básicos y características.

La representación gráfica se considera un potente mecanismo de transmisión de información y se ha convertido en una forma habitual de comunicar contenidos, tanto de carácter científico como divulgativo, en prensa especializada, Internet y medios audiovisuales. La creación y difusión de gráficos e infografías se ha extendido y democratizado gracias al acceso a potentes equipos informáticos y a la red, lo que a su vez, también favorece un crecimiento inusual de nuevos datos.

Las disciplinas que se ocupan del diseño de la información consideran que los datos son la materia prima, pero carecen de valor informativo por sí mismos y, solo si están organizados, se convierten en información que puede ser comprendida, memorizada y transformada en conocimiento por el usuario (Vid. Fig. 2).

El *diseño de la información* comprende el conjunto de técnicas y herramientas que preparan la información para que pueda ser usada fácilmente por los lectores, a menudo fatigados y sobrecargados con la abundancia de estímulos informativos

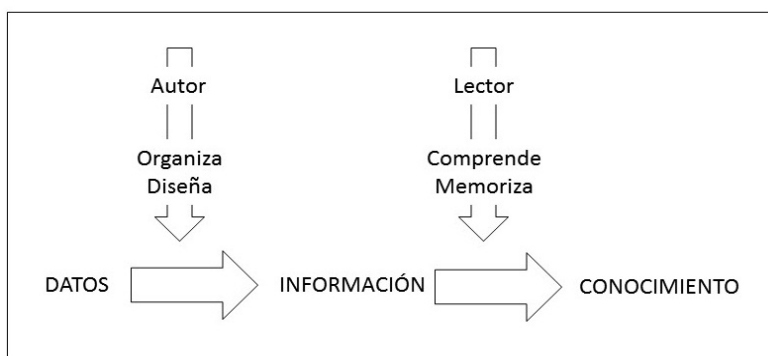



Fig. 2. La cadena de los datos, la información y el conocimiento
Fuente: elaboración propia

Dentro de esta disciplina, bebiendo también de las técnicas de la comunicación gráfica, se enmarca la *visualización de la información*, que

de forma específica se ocupa de la presentación diagramática de los datos y de su transformación visual para facilitar la comprensión de los mismos. Comprende dos áreas interrelacionadas: el modelado estructural y su representación gráfica y define tanto los procesos de análisis y preparación de los datos como la representación mediante el constructo gráfico. El diseñador de visualizaciones identifica, selecciona relaciones y patrones en los conjuntos de datos y posteriormente los representa gráficamente sobre un soporte (CAIRO, 2008, p. 27-28).

Se puede definir la visualización de la información como una disciplina transversal que "utiliza el inmenso poder de comunicación de las imágenes para explicar de manera comprensible las relaciones de significado, causa y dependencia que se pueden encontrar entre las grandes masas abstractas de información que generan los procesos científicos y sociales" (MEDIALAB PRADO, 2016), puesto que "se encarga de la representación visual de contenidos proposicionales mediante el uso de diagramas, gráficas y esquemas para facilitar la aprehensión, la interpretación, la transformación y la comunicación de esos contenidos a través de esas representaciones visuales" (PÉREZ-MONTORO, 2009).

El diseño y la visualización de la información son conceptos estrechamente ligados al más genérico de *arquitectura de la información*, utilizado por primera vez por Richard Saul Wurman en 1975, para designar el estudio de la organización de los datos con el objetivo de permitir al usuario encontrar su vía de navegación hacia el conocimiento y la comprensión de la información.

Wurman, arquitecto de profesión, estaba interesado en la clase de interacción que se producía entre las personas y su ámbito urbano y en el tipo de medios que podían ayudar a transmitir la información de estos entornos a los profesionales de la arquitectura, ingenieros, turistas y a los ciudadanos en general. La suya era una concepción más próxima al mundo del diseño gráfico y a la visualización de información. Pero, unos años más tarde, en1998, Louis Rosenfeld y Peter Morville adoptaron el término extrapolándolo al ámbito del diseño de sitios web y sistematizaron por primera vez los principios de la emergente disciplina.

La arquitectura de la información es un proceso interactivo y transversal, que no busca definir una metodología de diseño universal sino articular un conjunto de técnicas para ayudar al desarrollo y producción de espacios de información como los sitios web.

Los tres conceptos –arquitectura, diseño y visualización de la información– son partes específicas, intrínsecamente relacionadas con el marco global que proporciona la denominada *experiencia de usuario*, en la que intervienen tanto la arquitectura de la información como el diseño de interacción, la usabilidad, la accesibilidad, el diseño gráfico, la estética y la psicología cognitiva, entre otras disciplinas (Vid. Fig. 3).

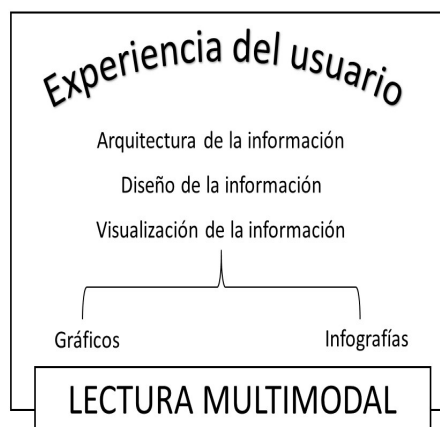



Fig. 3. Interrelaciones entre la arquitectura, el diseño y la visualización de la información
Fuente: elaboración propia

El conjunto de factores que determinan la interacción satisfactoria del usuario con un entorno o dispositivo concreto, siendo capaces de generar en él un conjunto de percepciones y emociones positivas sobre el medio y su uso, conforman lo que se conoce justamente con esa denominación ya mencionada: *experiencia de usuario*. Esta disciplina, como la arquitectura y la visualización de la información se articula en torno a los conceptos clave de *diseño de la información, percepción e interactividad*. La madurez de dicha disciplina hace que hoy sea una tautología decir lo que había que afirmar y justificar hace años: que el diseño (también) comunica.

3. Unas breves pinceladas históricas sobre la comunicación gráfica

A lo largo del tiempo el ser humano ha sido capaz de plasmar de diversas formas gráficas determinada información científica, cultural o de cualquier otra naturaleza. Es el caso de disciplinas como la cartografía –con su "naturaleza abstracta, simbólica y relacional" (CAIRO, 2008, p. 39)–, la estadística o la ilustración científica propiamente dicha, mediante mapas, tablas y gráficos y dibujos científicos, respectivamente.

Se considera que los mapas, que muestran esquemáticamente relaciones entre elementos geográficos, son las representaciones visuales no figurativas más antiguas. Se conservan ejemplos pertenecientes al antiguo Egipto y China.

Podemos citar, por poner un ejemplo ya del siglo XX, el famoso plano del Metro de Londres, obra del ingeniero Henry Beck, considerado uno de los mejores diseñadores de información del mundo contemporáneo. Su creación se publicó en 1933, estaba inspirada en los diagramas de la red eléctrica y su éxito consistió en obviar las distancias reales entre las estaciones, utilizar pocos símbolos y colores distintos para cada línea. En un principio, la empresa del metropolitano fue escéptica frente a la radical propuesta de Beck, pero tuvo una rápida aceptación por el público, que lo encontró útil y simple. Su propuesta triunfó y determinó la forma de diseñar los planos de todos los suburbanos del mundo.

Fig. 4. Metro de Londres, diseñado por Henry Beck (1933).

También las primeras infografías en la prensa abrirían un camino evolutivo que alcanza hoy cotas de indudable calidad e incluso brillantez técnica y estética. El caso de los periódicos lo trataremos más ampliamente en el siguiente apartado. Digamos tan sólo que capítulo aparte habría que dedicar a la sección de gráficos del *The New York Times*, fundado en 1851 y que todavía hoy está considerado *la biblia* del periodismo, el diario más prestigioso del mundo.

Todos los casos expuestos son muestras históricas de la visualización de información creadas a lo largo del recorrido de los siglos. Han sido estudiadas con amplitud y precisión por Alberto Cairo (2008 y 2011), especialista de referencia que proporciona en su trabajos un excelente acercamiento a cada uno de estos tipos de ilustración.

4. Gráficos para la comunicación científica: características, tipos y uso para la comunicación y la divulgación de la ciencia

Los gráficos son representaciones visuales que enriquecen la información proporcionada por cualquier tipo de estudio o investigación, al aportar diversas fuentes semióticas –verbales, icónicas, numéricas– que contribuyen a comunicar el significado de los textos científicos de forma más condensada y eficaz. Se caracterizan por atraer y captar la atención del lector presentando los datos y la información de forma clara y precisa.

Los gráficos se pueden clasificar en *gráficos de tipo estadístico* y *gráficos de tipo no estadístico*.

Los gráficos de tipo estadístico son las más utilizadas por la comunidad científica para comunicar y difundir los resultados de las investigaciones realizadas. Dentro de ellos, existen distintos tipos, según se busque comparar, detectar mayorías o minorías, determinar tendencias, incidencias, evoluciones, etc.

Los gráficos estadísticos más usuales son: el diagrama de barras;el gráfico o diagrama de sectores; el histograma y el polígono de frecuencias.

A) El diagrama de barras se utiliza para presentar datos cualitativos o datos cuantitativos de tipo discreto, comparando dos o más valores. Los datos se representan mediante barras de una altura proporcional

a la frecuencia, sobre unos ejes de coordenadas o ejes cartesianos. Las barras pueden tener una orientación vertical u horizontal.

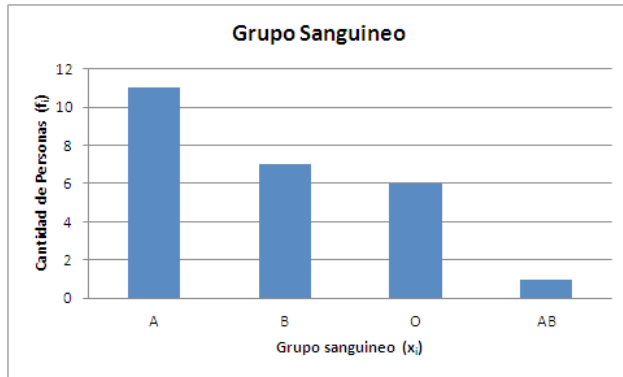



Fig. 5. Diagrama de barras.

B) El gráfico o diagrama de sectores es una representación circular de las frecuencias relativas de una variable cualitativa o discreta que permite, de una manera sencilla y rápida, su comparación. El círculo representa la totalidad que se quiere observar y cada porción, llamada sector, representa la proporción de cada categoría de la variable respecto el total. Suele expresarse en porcentajes. Son útiles cuando las categorías son pocas. Si el gráfico tuviera muchas variables, no aportaría casi información y sería prácticamente incomprensible.

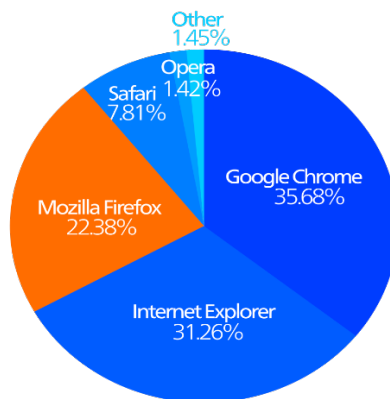



Fig. 6. Diagrama de sectores.

C) El histograma se usa para representar las frecuencias de una variable cuantitativa continua.En uno de los ejes se posicionan las clases de la variable continua y en el otro eje las frecuencias. No existe separación entre las barras.

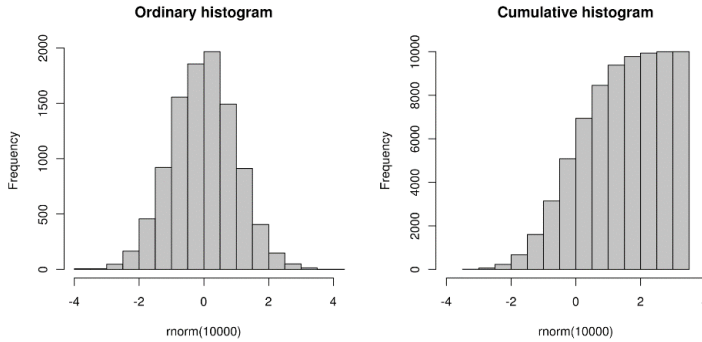



Fig. 7. Tipos de histogramas.

D) El polígono de frecuencias es un tipo de gráfico que se crea a partir de un histograma de frecuencia. El polígono surge de unir los puntos medios de las bases superiores de las barras de un diagrama de barras, e incluso también de un histograma.

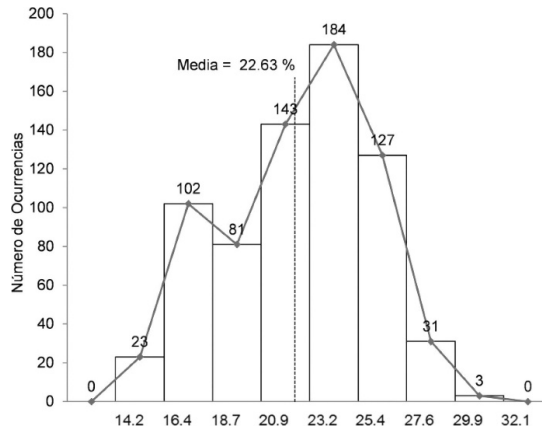



Fig. 8. Polígono de frecuencias.

Entre los principales gráficos de tipo no estadístico destacan los pictogramas y los cartogramas

A) Los pictogramas son gráficos en los que se usan signos que representan de forma figurativa, más o menos realista, elocuente y esquemática, objetos o ideas. Se encuentran en los orígenes de la mayoría de los sistemas de escritura.

Fig. 9. Ejemplo de pictograma.

B) Los cartogramas son esquemas estadísticos y cartográficos de una superficie, que muestran el tamaño de esa superficie en función del determinado valor estadístico que ella represente, siendo ese valor más importante de cuantificar que el valor de la superficie. Su objetivo es ilustrar de forma impactante una distribución temática sobre un territorio considerando que la superficie de las unidades territoriales adquiere un tamaño proporcional al valor de la variable representada.

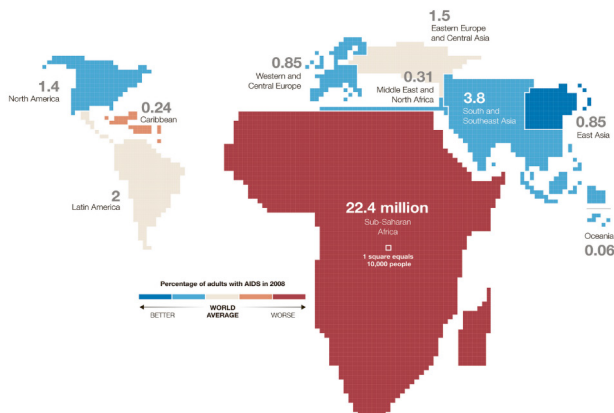



Fig. 10. Ejemplo de cartograma.

Los modelos de gráficos presentados integran fuentes semánticas variadas – verbales, icónicas y numéricas– caracterizadas por una alta condensación temática. Requieren por tanto, una lectura multimodal que interrelacione los significados lingüísticos, visuales y matemáticos para construir la coherencia global de los textos (GUZMÁN SIMÓN, 2016, p. 26-27).

5. Infografías para la información de actualidad: características, tipos y usos informativos y periodísticos.

La infografía es "una representación diagramática de datos". En tanto que un diagrama es "una representación abstracta de una realidad". Se trata de un acrónimo derivado del término anglosajón *infographics* y éste de la expresión *information graphics*. Las infografías se popularizan como consecuencia de la informatización de las redacciones periodísticas, a fines de los años ochenta y principios de los noventa (CAIRO, 2008, p. 21).

Ha sido definida por algunos autores, en los albores de su abordaje académico dentro de esa etapa, como un nuevo género periodístico (PABLOS, 1991, p. 153), una concepción reiterada después por otros investigadores.

Más allá de esa catalogación teórica, en cierta medida discutible, no es menos cierto que en realidad se trata de una forma de expresión que adopta enfoques novedosos, y se muestra en auge progresivo por la sofisticación de la tecnología y el diseño gráfico e interactivo, que suele estar al servicio de los grandes géneros periodísticos clásicos.

De este modo, un reportaje en profundidad o un informe o dossier especial (como el que mostraremos después sobre el tren AVE español) ganarán mucho con la presencia de infografías. Estas raras veces se presentan aisladas, sino como un elemento que forma parte de un conjunto que le da solidez y lógica informativa, al ser complementarios texto e imagen. Eso no significa que carezcan de la autonomía y el valor informativo y documental intrínsecos a la infografía en sus diversas tipologías de plasmación, que serán expuestas más adelante.

De hecho, la hoy conocida como primera Guerra del Golfo (1990-91) supuso un antes y un después en la cantidad y la calidad de infografías publicadas por la prensa. Se elaboraron unas piezas cuya carga semántica –desde el punto de vista del periodismo y la documentación, pero también en su dimensión estética o artística– proporcionaba mucho más que un brillante apoyo a las crónicas de guerra de los reporteros y enviados especiales.

En el caso español, los periódicos ofrecieron un trabajo de gran calidad, a la altura de la prensa europea de referencia, que también vivió una eclosión en este terreno, en sintonía con los recursos desplegados para contar al mundo la considerada como primera guerra televisada en directo, en concreto por la indiscutible supremacía en aquella época de la cadena norteamericana CNN.

Hoy se ha producido un salto cualitativo mediante la generalización del color en los grafismos de la prensa, también de la impresa, frente al blanco y negro de los primeros años noventa en esas infografías bélicas a las que nos hemos referido.

Los estudios sobre mensajes visuales constatan la relevancia de los grafismo tanto en los relatos publicitarios (MARTÍNEZ PASTOR, MONTES VOZMEDIANO, 2015) como en los periodísticos (MONTES VOZMEDIANO, 2015).

Es de nuevo Alberto Cairo (2008, p. 31) quien ha expuesto que los malentendidos son habituales incluso en la propia definición de lo que es infografía. Y repasa diversas formas de concebirla: Almeida la describe como “info + grafía = texto + imagen”, sugiriendo inadvertidamente que la imagen no informa, sino que es el texto el que aporta el necesario empaque para que la infografía pueda ser considerada producto periodístico.Por su parte, González Díez y Pérez Cuadrado (2001) la consideran como un tipo específico de ilustración y la denominan “ilustración informativa”.

Diferentes autores ofrecen formas muy disímiles de concebir la visualización de la información: como uso más o menos intercambiable con la fotografía (PABLOS, 1999) o como “simplemente un recurso gráfico” al servicio del reportaje (YANES MESA, 2004). También es considerada parte de una serie de “nuevos formatos gráficos” que permiten apelar a un

lector que cada vez con mayor frecuencia y asiduidad se encuentra frente a una pantalla (LÓPEZ HIDALGO, 2002).

Como expone Rosana Larraz (2016), las infografías, "son más sintéticas que los vídeos, más narrativas que un esquema, más atractivas que las tablas de datos, más exploratorias que las presentaciones tradicionales y, a diferencia de los textos escritos, permiten visualizar la información que presentan. Las infografías transmiten hechos, procesos, noticias, acontecimientos o datos de forma amena, sintética y visual, facilitando la comprensión de información árida o compleja y estimulando el interés del lector que, de un golpe de vista, puede seleccionar en ellas lo que le interesa, lo que ya conoce y lo que no. En formatos estáticos o animados, impresos o digitales integrando elementos multimedia y a menudo interactividad, las infografías ofrecen muy variadas posibilidades de explotación didáctica".

Continúa la autora: "pueden ser canteras de información y motores para la expresión oral y escrita en el aprendizaje de lenguas, facilitadoras de la comprensión de los contenidos curriculares, estímulos para el repaso o la ampliación, recursos que motiven el conocimiento de la actualidad o bien vehículos de expresión y creatividad utilizando las nuevas tecnologías para su producción. La existencia en Internet de gran cantidad de información narrada en este género añade a su utilización la oportunidad de contar con contenidos siempre actualizados. Trabajar con infografías en el aula supone, además, introducir en la enseñanza los nuevos modos de expresión visual y multimedia"[2].

En el caso de la infografía digital, presente en el ciberperiodismo, esta modalidad, "tiene capacidad de expresión periodística y documental con importantes desarrollos visuales que le son propios" (VALERO SANCHO, 2008):

1. En el contenido, presenta temas recurrentes y muy apropiados como las guerras, actos de terrorismo, elecciones, grandes catástrofes, documentales, acontecimientos deportivos y cualquier tipo de informaciones relevantes que suelen desplegarse en toda su extensión.

[2] Rosana Larraz, "Infografías como recursos didácticos" (2016). *Cuaderno intercultural*, disponible en http://www.cuadernointercultural.com/infografias-recurso-didactico/) [consulta 25 abril 2017].

2. En la forma de presentación y estructura, por medio de sus recursos de tipo digital e incluso multimediático.
3. Empleando herramientas características de los principales programas de edición y diseño web.

Como ejemplo de ello, puede consultarse el dossier especial sobre los 25 años del tren de Alta Velocidad (AVE) en España, publicado por la edición digital del diario *El País*(Vid. Fig. 11, 12 y 13).

En él se incluyen números gráficos e infografías, como la imagen simétrica que permite deslizar el cursor de un lado a otro para ver la fotografía en blanco y negro del primer tren que circuló en 1992 entre Madrid y Sevilla o la mostrada en color de un convoy actual. Ambas imágenes superpuestas permiten descubrir un tren más moderno que el que aparece en la fotografía que muestra un vehículo que a ojos de hoy parece obsoleto, por sus líneas menos dinámicas y su diseño menos avanzado.

Sirve de ejemplo sobre cómo se utilizan estos recursos, muy vistosos, particularmente en ocasiones que trascienden la información diaria, las noticias cotidianas. Se debe al esfuerzo que supone la elaboración de esos materiales, a cargo de profesionales especializados y de alta cualificación, algo que no todos los medios de comunicación pueden permitirse.

Suponen creaciones que actúan como elementos de jerarquización de la información, para conceder relevancia a determinados temas o acontecimientos. Esta función se ejercía, y se ejerce todavía, de forma muy evidente en la prensa convencional –la de papel– mediante el espacio dedicado (un *faldón* o cuarto inferior de página, media, una completa, dos…) y el lugar ocupado (en la parte superior o no de la página par o impar).

Sin embargo, en las publicaciones digitales, aunque lógicamente las noticias y reportajes tienen igualmente una extensión determinada, y aparecen en la parte superior, central o inferior del diario electrónico, el propio ritmo y lógica de la navegación hacen menos visibles y marcados esos elementos de la prensa tradicional, que en parte se compensan con la presencia de imágenes y el despliegue de recursos interactivos visuales.

Fig. 11. Fotografía en blanco y negro de 1992.

Fig. 12. Superposición de dos fotografías de épocas diferentes.

Fig. 13. Fotografía en color de 2017-2018.
Fuente: elpaís.com, 20 de abril de 2017:
http://economia.elpais.com/economia/2017/04/20/actualidad/1492675474_811055.html
[consulta 20 abril 2017].

De acuerdo con Valero Sancho, Catalá Domínguez y Marín Ochoa (2014), ofrecemos a continuación una tipología de visualizaciones o infografías. Tanto la clasificación como las definiciones son de los autores citados. A su trabajo teórico y análisis aplicado remitimos para ver sus ejemplos y sus explicaciones extensas y particulares, que incluyen aclaraciones sobre el procedimiento profesional, como cuando resulta necesaria la activación, algoritmos y datos introducidos en el documento para un determinado cálculo.

Los distintos tipos de visualizaciones de la información son los siguientes:

1. Visualización espacial: es la comparación de espacios, confluencias de puntos, líneas o zonas coloreadas a partir de datos.
2. Visualización tabular: son estudios de propiedades registrales, interrelaciones comparadas, relaciones entre registros o campos, etc.
3. Visualización posicional: la posición en el espacio y el tipo de lenguaje empleado según la importancia que tengan puede denotar significación o relaciones de influencia.
4. Visualización topográfica: la representación por medio de territorios (mapas, planos, recintos o microplanos) puede o no tener objetivos ubicativos y también puede ser un soporte para la distribución de datos o figuras, con diversas propiedades multiescalares, anamórficas, multicolores, interactivas, etc.
5. Visualización teledinámica: por la importancia que tienen,los autores mencionados distinguen las presentaciones dinámicas cambiantes, procesos de automatismos, entre otros.
6. Visualización de arrastre interactivo (*Drag and drop*): se puede considerar la más interactiva en general. Se presenta con funciones de arrastre o elección por parte del intérprete que activa, cálculos o proposiciones en diagramas de resolución, calculadoras, etc.
7. Visualización de identificación aumentada: se trata de evoluciones gráficas de un proceso a partir de datos e imágenes fotográficas de alta resolución de escenas captadas y datos asociados, que aparecen al solicitarlas el intérprete.

8. Miscelánea de varias visualizaciones: son multitud de posibles representaciones conjuntas de gráficos, simultáneas o sucesivas, que se pueden presentar mezcladas plurimórficas y multipropuesta, como complementos, etc.

6. Conclusiones

Hoy es un lugar común aseverar que el diseño gráfico *también* comunica. Sin embargo, no hace tanto tiempo se infravaloraba en muchos ámbitos, incluido el campo del periodismo y la comunicación profesional, todo lo que tuviera que ver con la imagen y con sus posibilidades frente al texto, ya fuese fotografía o cualquier otro elemento visual o gráfico. Las posibilidades que brinda el desarrollo tecnológico, junto con la progresiva generalización y aceptación y uso social, han acabado con esa visión y esos tratamientos injustificados.

Si la fotografía en el ámbito informativo-periodístico se ha convertido –tanto en su propio presente como, especialmente, al pasar del tiempo– en un auténtico documento social, reflejo de épocas, situaciones, personajes y ambientes (TORREGROSA CARMONA, 2010), esa realidad se comparte hoy con la visualización de la información mediante infografías que explican hechos del presente.

En este sentido, se cumple a la perfección en el universo significativo de la imagen, fotoperiodística o infográfica, aquella verdad de la documentación según la cual a medida que disminuye el valor inmediato o incluso administrativo de un documento de cualquier clase, aumenta su valor patrimonial o histórico, al dar testimonio de valores étnicos, antropológicos, ideológicos, sociales, culturales, etc.

En los momentos actuales, tanto los gráficos como las infografías juegan claramente a favor de la eficacia comunicativa, la comprensión y la contextualización de las informaciones proporcionadas.

Los periódicos lo están sabiendo aprovechar, en parte gracias a un lector más acostumbrado a ello y mucho más exigente con la visualización de la información, y en parte gracias a profesionales especializados y bien formados.

Unos especialistas que conocen y sacan partido a las nuevas aplicaciones y posibilidades tecnológicas en este campo en expansión que sigue explorando lenguajes y códigos icónicos. De alguna manera, esa audiencia *juega* con la nueva forma de exposición y comunicación de la información, recreándola y renovando el pacto entre creadores y consumidores de la información.

Una nueva oferta y demanda en la visualización informativa que constituye una realidad creciente en los medios de comunicación, que cabe enmarcar dentro de la denominada *experiencia del usuario*. Y que contribuye, en un proceso bidireccional de causa y consecuencia, al llamado *periodismo de datos* gracias a los gráficos interactivos, los diagramas estadísticos y otros elementos visuales similares que apoyan los trabajos más relevantes de investigación periodística.

Esa denominación de *periodismo de datos*, tan en boga ahora, es comparable a la de periodismo de precisión, que estuvo en vigor decenios atrás. En realidad, son parcelas parecidas, que otros autores engloban en las actividades generales de la investigación.

Y es que tanto la información como la documentación disponen con estas creaciones profesionales de un recurso valioso para suscitar el interés del lector y promover la mejor comprensión tanto de las investigaciones científicas, como de las informaciones de actualidad.

7. Bibliografía

ALCALDE, I. *Visualización de la información*. Barcelona: UOC, 2015.

AYUNTAMIENTO de Madrid Visualizar. *Medialab-Prado*, 2016 [consulta 23 abril 2017]. Disponible en: http://medialab-prado.es/visualizar.

CAIRO, A. *Infografía 2.0.*Vistualización interactiva de información en prensa. Madrid: Alamut, 2008.

CAIRO, A. *El arte funcional.* Infografía y visualización de la información. Madrid: Alamut, 2011.

GONZÁLEZ DÍEZ, L.; PÉREZ CUADRADO, P. *Principios básicos sobre diseño periodístico*. Madrid: Universitas, 2001.

GONZÁLEZ, J. A.; JOVER, L. El impacto de los gráficos en la comunicación científica. In: JIMÉNEZ VILA, J. et al. *Publicación científica biomédica.* 2. ed. Barcelona: Elsevier, 2015.

GUZMÁN SIMÓN, F. La alfabetización multimodal en la educación superior. In: GÓMEZ CAMACHO, A. (Coord.) *La alfabetización multimodal:* nuevas formas de leer y de escribir en el entorno digital. Madrid: Síntesis. Psicología/ Educción, 2016. p. 17- 32.

LANKOW, J.; RITCHIE, J.; CROOKS, R. *Infografías.* El poder del storytelling visual. Barcelona: Gestión 2000, 2013.

LARRAZ, R. Infografías como recursos didácticos. *Cuaderno intercultural:* Recursos para la interculturalidad y la educación intercultural, 2016. [consulta 25 abril 2017]. Disponible en http://www.cuadernointercultural.com/ infografias-recurso-didactico/.

LÓPEZ HIDALGO, A. *Géneros periodísticos complementarios:* una aproximación crítica a los formatos del periodismo visual. Sevilla; Zamora: Comunicación Social; Ediciones y Publicaciones, 2002.

MARTÍNEZ PASTOR, E.; MONTES VOZMEDIANO, M. Pain in advertising: an analysis of the content and design of the advertisements in press after 11-M terrorist attacks. *Communication & Society*, v. 28, n. 1, p. 115-126, 2015.

MONTES VOZMEDIANO, M. Las funcionalidades de los grafismos en los mensajes periodísticos: análisis y propuesta sistematizadora. *Estudios sobre el Mensaje Periodístico,* v. 21, n. 1, p. 421-435, 2015.

PABLOS, J. M. La infografía, el nuevo género periodístico. In: BARTOLOMÉ CRESPO, D. et al. *Estudios sobre tecnologías dela información*. Madrid: Sanz y Torres, 1991. p. 152-190.

PABLOS, J. M. *Infoperiodismo:* el periodista como creador de infografía. Madrid: Síntesis, 1999.

PÉREZ-MONTORO, M. Visualización de la información || Information Visualization. *Glossarium-BITri, 2009.* [consulta 20 abril 2017]. Disponible en: http://glossarium.bitrum.unileon.es/Home/visualizacion-de-la-informacion

TORREGROSA CARMONA, J. F. Modelos para el análisis documental de la fotografía. *Documentación de las Ciencias de la Información*, v. 33, p. 329-342, 2010. [consulta 27 abril 2017]. Disponible en: http://revistas.ucm.es/index.php/DCIN/article/view/19656.

VALERO SANCHO, J. L. La infografía digital en el ciberperiodismo. *Revista Latina de Comunicación Social,* v. 63, p. 492-504, 2008. [consulta 20 abril 2017]. Disponible en: http://www.ull.es/publicaciones/latina/08/42_799_65_Bellaterra/Jose_Luis_Valero.html.

VALERO SANCHO, J. L.; CATALÀ DOMÍNGUEZ, J.; MARÍN OCHOA, B. E. Aproximación a una taxonomía de la visualización de datos. *Revista Latina de Comunicación Social*, v. 69, p. 486-507, 2014. [consulta 27 abril 2017]. Disponible en: http://www.revistalatinacs.org/069/paper/1021_UAB/24es.html.

VALERO SANCHO, J. L. *La infografía.* Técnicas, análisis y usos periodísticos. Valencia: Universitat de Valencia; Castelló de la Plana: Publicacions de la Universitat Jaume I; Barcelona: Universitat Pompeu i Fabra; Bellaterra: Universitat Autònoma de Barcelona. Aldea Global; 9, 2001.

YANES MESA, R. *Géneros periodísticos y géneros anexos*. Madrid: Fragua, 2004.

A SEMIÓTICA DOCUMENTAL COMO UM ESPAÇO INTERDISCIPLINAR PARA O ESTUDO DA LEITURA EM INDEXAÇÃO

Carlos Cândido de Almeida

1 Introdução

A questão que nos cabe discutir neste trabalho está intimamente ligada ao processo semiótico denominado de indexação de assunto. Recorremos a definição de Lancaster (1993) para quem a indexação corresponde às atividades de representação do conteúdo de partes de um documento com o objetivo de produzir índices que levem ao assunto de uma obra. Sabemos que esta linha se relaciona à catalogação de assunto e à classificação. Para Lancaster, a etapa de indexação tem como processos a análise conceitual e a tradução. Há um ligeiro contraste entre esta perspectiva e o entendimento de indexação dos autores ligados à matriz francesa da análise documental (GARDIN, 1966; CHAUMIER, 1971; CUNHA, 1989). Para estes últimos, a indexação é uma fase que tem como etapa anterior à segmentação do conteúdo dos documentos (extração, análise, fase analítica ou leitura). Nesta etapa anterior ao processamento dos documentos que se instaura o problema da leitura documental, isto é, a leitura voltada aos interesses da representação e, consequentemente, da recuperação da informação (FUJITA, 2004).

Não estamos interessados em contrapor ou endossar argumentos sobre a querela terminológica entre indexação de assunto e análise documental francesa, se é de fato que ela existe. A questão que nos interessa responder é qual disciplina, campo, área, núcleo de pesquisa ou ciência poderia dar o suporte devido e abrigar as diversas contribuições para compreender o processo de leitura documental. Em outras palavras, que nicho a leitura documental pode supor como sua estrutura teórica. Primeiramente, devemos pensar a leitura documental de uma maneira pouco usual que a vê como um processo mais amplo, neste caso, o processo de leitura profissional não se restringiria à decomposição de conceitos de livros e identificação do assunto principal para a representação de textos escritos. Em outras palavras, a leitura documental, como o próprio nome sugere, analisaria a linguagem de toda sorte de documentos, de uma variedade de códigos. Em segundo lugar, acreditamos que a base teórica que daria suporte à leitura documental não poderia se restringir a uma Linguística que lhe é própria, tal como concebida nos anos 1980 por García Gutiérrez.

Entendemos que a Linguística Documental, sob o ponto de vista do impacto que teve em seu país de origem, foi um projeto de disciplina associada à documentação, tal como esta se configurava na Espanha no início dos anos 1980. Chamamos projeto em virtude da tentativa malograda de impulsionar a criação de disciplinas de Linguística Documental nos currículos dos cursos de Diplomatura e Licenciatura em Biblioteconomia e Documentação. O que se impôs nestas tentativas foi a já conhecida disciplina Linguagens Documentais nos cursos para formação de documentalistas na Espanha.

O processo de leitura documental, como uma das fases do processo geral de indexação, deve requerer um campo teórico específico que não restrinja o número de variáveis envolvidas. Para tanto, devemos cogitar a possibilidade de realocar a leitura documental de campo disciplinar.

Desse modo, o objetivo aqui é explicar a importância epistemológica de uma Semiótica Documental que acrescentasse, em extensão, elementos conceituais à leitura documental. Com esse objetivo em vista, devemos apresentar sumariamente a Linguística Documental e como

procura recobrir a área de interesse da Documentação[1], para, daí então, apresentar a ideia básica da Semiótica Documental, a qual pode abrigar plenamente as variáveis do processo de leitura documental.

2 A Linguística Documental: espaço para refletir sobre a leitura documental?

A Linguística Documental é um intento disciplinar fundamental para compreender as linguagens documentais na Documentação espanhola. Com alguma reserva utilizamos a palavra "disciplina" pois não nos parece consensual a sua posição diante da Documentação, e principalmente da Linguística. Contudo, na versão espanhola da Linguística Documental – muito utilizada por autores brasileiros (LARA; TÁLAMO, 2006, 2007) –, as questões ligadas à leitura e à indexação de assunto ficam em segundo plano ou são totalmente desconsideradas em comparação a atenção colocada às linguagens documentais.

Uma das obras mais influentes sobre o assunto foi publicada por García Gutiérrez, em 1984. Na oportunidade propôs a Linguística Documental como uma "interdisciplina de lenguajes documentales", ou melhor, "[…] una disciplina impregnada de outros campos científicos, como son basicamente la Lógica, la Estadística, y la Informática y más concretamente, la Lexicología, la Archivística, la Biblioteconomía y la Telemática, bajo los imperativos específicos de ámbito del saber que va a ser controlado por un lenguaje documental." (GARCÍA GUTIÉRREZ, 1984, p.138).

As bases teóricas gerais da Linguística documental, arroladas por García Gutiérrez (1990) em outro trabalho, são provenientes da Linguística Geral, da Semântica, da Linguística Textual, da Sociolinguística, do Estruturalismo, do Gerativismo, da Teoria da Comunicação e da Informação, da Teoria do Contexto, da Teoria da representação e da Teoria da Tradução. Sem perder a sua individualidade e pretensa autonomia diante destes campos, a Linguística Documental recorre a estas abordagens para cobrir aspectos de interesse à Documentação. A Semiologia e a Semiótica aparecem como disciplinas indiretamente relacionadas à Documentação,

[1] Utilizamos neste trabalho o termo "Documentação" como equivalente à "Ciência da Informação".

ora como contexto geral à descrição da Linguística Saussureana (GARCÍA GUTIÉRREZ, 1990, p. 44), ora como teoria vinculada à Semântica, Comunicação e Teoria da Representação. É sugestiva a seguinte citação: "Sin embargo, el documentólogo no está especialmente interesado por las funciones simbólicas (terreno de lingüistas y semiólogos) o por las funciones simbólicas (terreno de psicólogos)." (GARCÍA GUTIÉRREZ, 1990, p. 46).

Esta disciplina vale-se da linguagem documental que, segundo o autor, é seu principal objeto. Por linguagem documental entende um sistema de signos naturais ou artificiais para a identificação dos documentos, empregados na análise e no recuperação dos documentos. Essa linguagem funciona como meio de expressão criado para, de um lado, o controle, por outro, para a comunicação, isto é, uma função intermediária.

O autor põe em evidência uma acepção de linguagem. Nesse sentido, linguagem seria um sistema de signos inscrito em um documento. "El sistema de signos escritos en el documento es lenguaje humano pero el sistema de signos naturales o artificiales utilizados para identificar ese documento, entre un conjunto de ellos, es lenguaje documental, aunque éste no pueda existir si el documento carece del contenido que le otorga la acción intelectual humana." (GARCÍA GUTIÉRREZ, 1984,p. 136).

A linguagem é um sistema de signos, fato que contrasta com a definição estruturalista de linguagem que se aduze em Barthes e Saussure. O dispositivo de execução da linguagem – fala ou outro mecanismo de expressão - está ausente nesta definição de García Gutiérrez. Devemos sublinhar ainda que esta compreensão de linguagem que focaliza o sistema e o código, tem sido amplamente aceita na literatura da Documentação, dentro e fora da Espanha.

O conceito de Linguística empregado pelo autor é sintetizado pela famigerada fórmula: ciência que estuda a linguagem humana. De um ponto de vista aplicado, García Gutiérrez (1984, p.139) enumera os conceitos derivados ou relativos ao cruzamento disciplinar com a Linguística, a saber: *descritor*, ou unidade significativa mínima da mensagem documental; *frase documental*, que entendemos por conjunto de descritores ligados ou não por relações sintáticas artificiais e *resumo documental*, um produto

resultante da aglutinação de mensagens de um documento que são expressáveis em linguagem controlada.

Estas novas aplicações, concretamente novas acepções conceituais, parecem ser resultado da comparação teórica entre as disciplinas em questão, utilizando como parâmetro os termos especializados da Linguística para explicar fenômenos afeitos à Documentação. De uma maneira mais específica, esta explicação resolveria a lacuna da definição de linguagem que se tem adotado, pois não é apenas um sistema de signos, mas um código que pode ser combinado e executado, mesmo que não seja por obra de um falante em "documentês".

A proposta de García Gutiérrez ampara-se em uma linha semiótica distinta da proposta de Charles Peirce (1839-1914), ainda que se distancie de uma superficial aproximação à Linguística, pois seu trabalho é criterioso em termos de armação conceitual. Contudo, a leitura, a qual poderia ser admitida como processo de decomposição da linguagem dos documentos para a representação da informação no contexto documental, ou seja, não é elencada como objeto principal de análise. Nesse sentido, a tentativa de instauração da Linguística documental, que poderia significar a verdadeira confluência entre Linguística e Documentação nos diversos temas de interface – entre os quais a leitura – negligenciou objetos candentes das relações disciplinares entre os campos citados.

A despeito da escolha da Linguística Documental pela linguagem documental, há importantes questões associadas ao processo de leitura para a indexação, como foi o caso do debate sobre linguagem científica e técnica que é entendido pelo autor (GARCÍA GUTIÉRREZ, 1984, p. 140) como um léxico especializado, empregado por uma comunidade em uma disciplina ou setor científico. Como a ciência depende da comunicação entre várias comunidades científicas, o documentalista, assim como linguistas e terminólogos, seriam os guardiães da língua (GARCÍA GUTIÉRREZ, 1984, p.141). Essa linguagem é um léxico, pois não é diferente da língua dominante, uma vez que extrai dela a estrutura, a despeito do uso de um distinto vocabulário (GARCÍA GUTIÉRREZ, 1984, p.140).

O autor defende que as linguagens documentais não podem ser apenas vocabulários técnicos, como assevera Trujillo (apud GARCÍA GUTIÉRREZ, 1984, p. 144).

> Este razonamiento no es válido con el de los lenguajes documentales, en los que, si bien a niveles simples, se dan relaciones semánticas entre los conceptos e incluso sintácticos, por tanto hay que hablar de lenguajes dinámicos frente al estatismo interno de un vocabulário. (GARCÍA GUTIÉRREZ, 1984, p.144)

O autor defende a tríade documentação-terminologia-normalização. O que procura argumentar com a afirmação é que o documentalista deve ter conhecimentos de terminologia e normalização (GARCÍA GUTIÉRREZ, 1984, p. 147). Assim, as teses gerais da Linguística documental a colocam em um espaço de encontro entre disciplinas, cujo propósito nem sempre foi contemplado pelas tradicionais abordagens da Documentação.

De acordo com García Gutiérrez (1990), o processo documental excede a natureza lógica e mecanicista, pois possui elementos linguísticos. Além disso, a Linguística documental não deve ser vista como uma Linguística aplicada, mas como uma genuína teoria da documentação cujo método está centrado na noção de estrutura e sistema relacional, para compreender a estrutura da documentação, enquanto conjunto de dados e informações sobre um tema. Essa natureza epistemológica da Linguística documental que precisa ser aclarada, isto é, saber até que ponto se está sendo processado um salto qualitativo na proposição de teorias ou simplesmente se aplica termos e conceitos que apenas sofisticam a argumentação e a defesa da Documentação. No constante à estrutura da documentação, é sugestiva a citação:

> La estructura de la Documentación, como organización de contenidos codificables y decodificables, es el objeto de la Lingüística documental en un doble sentido: 1) La estructura de la producción de información, la formación, organización y presentación de las ideas por parte del productor, como corpus de observación y descripción; 2) La estructura de la representación del discurso del productor, a la que se accede mediante formulaciones metodológicas y modelos de sínteses y traducción." (GARCÍA GUTIÉRREZ, 1990, p. 24).

Em teoria, tanto a análise documental de conteúdo quanto a linguagem documental deveriam ser assuntos pertinentes à Linguística documental como disciplina nos currículos dos cursos de Documentação na Espanha. Nesta segunda versão de Linguística Documental, podemos identificar o interesse pelo processo de análise, o qual contem a leitura documental para etapa geral de decomposição de conceitos. No entanto, exclusivamente no quesito oferta acadêmica e expansão desse entendimento junto ao circuito universitário, a Linguística documental revelou-se como um fato isolado na Espanha.

A questão decorrente é que a Linguística documental não forjou uma proposta clara e interdisciplinar para o entendimento do processo de leitura para indexação. Neste caso, supomos que se assemelha à noção de leitura (análise) tal como enseja a proposta de análise documental de Chaumier (1971) e Gardin (1974).

Outra dificuldade não tão contemporânea assim, é a insistência pelas propostas de análise de textos (códigos verbais) como os principais representantes do pensamento científico e especializado. Esta preferência pode ter feito ruir a base de uma acepção de leitura dedicada à compreensão dos mais distintos códigos não verbais, os quais constam em documentos que devem ser processados e representados para a recuperação.

Em suma, a Linguística documental, tal como foi concebida em sua origem não poderia abrigar uma proposta interdisciplinar do estudo da leitura para indexação de textos verbais e não verbais. Entre os avanços da Linguística Documental neste quesito, podemos citar a compreensão da estrutura textual, uma de suas principais variáveis (FUJITA, 2004). A seguir, apresentaremos a hipótese que temos desenvolvido (ALMEIDA, 2011; ALMEIDA; GARCÍA MARCO, 2015) de que a organização da informação, em geral, tal como a leitura para indexação,em específico, mereceria um espaço interdisciplinar que focalize também outros códigos.

3 Semiótica Documental: um campo interdisciplinar para o estudo da leitura documental

O argumento aqui defendido é bem simples e objetivo. Se a Linguística Documental não pode avançar como espaço epistemológico

a abrigar as reflexões sobre leitura documental, qual área comportaria tal interesse no interior da Documentação. Podemos responder prematuramente que a análise documental, tal como importada da matriz francesa da Documentação pelo Grupo Temma no Brasil, já se preocupava com o assunto. Contudo, há um interrogante, a noção de leitura ensejada pode não ser suficiente para abarcar a diversidade de experiências de leitura profissional executadas fora do código verbal. Neste caso, tanto a Linguística documental, como visto, quanto a análise documental, não seriam suficientes para dar suporte epistemológico à leitura documental como estudada na atualidade (FUJITA, 2004; AGUSTÍN LACRUZ, 2006, 2015; MANINI, 2001, 2004; PATO 2014). O objeto desta seção é perspectivar a Semiótica Documental e relacionar sua concepção como necessária para a ampliação do escopo de estudo da Documentação, de modo a incluir uma concepção mais geral de leitura para fins de indexação.

No caso da Documentação na Espanha, a Semiótica nem sempre teve um potencial de articulação conceitual, tampouco foi influente para constituir-se uma corrente teórica coerente. Os trabalhos no sentido de uma Semiótica aplicada à Documentação ou Semiótica Documental surgiram na década de 1980. Contudo, em que pese a influência conceitual de Peirce em muitos trabalhos (HJORLAND, 2003; RABER; BUDD, 2003; ABREU; MONTEIRO, 2010; MOURA, 2006, 2007, 2011; MOURA; SILVA; AMORIM, 2002; LARA, 1993, 1999, 2003, 2006; FRIEDMAN; THELLEFSEN, 2011; THELLEFSEN, 2002, 2003, 2004; THELLEFSEN; THELLEFSEN, 2004; MAI, 1997a, 1997b, 2000, 2001), não foi oferecido uma proposta disciplinar que congregasse Semiótica e Documentação, tal como a Semiótica Documental.

Consideramos como as ideias de maior destaque nessa direção foram as de Izquierdo Arroyo, que cunhou a expressão "Semiótica Documental", utilizada desde final dos anos 1980. Em uma publicação de 1992, Izquierdo Arroyo destacou as pesquisas que então desenvolvia e cujo teor já indicava a preocupação com a estruturação da Semiótica Documental. Izquierdo Arroyo (1992) procurou esquematizar uma agenda de pesquisas em quatro grandes eixos: a) para uma teoria da representação documental; b) ensaio histórico de Semiótica Documental; c) a ciência da busca secundária; d) concepção lógico-linguística da Documentação.

Os dois últimos eixos foram iniciados e contam com publicações na literatura especializada. Mas são os dois primeiros que revelam a linha de análise e os fundamentos da Semiótica Documental e com a Semiótica de Peirce[2]. Izquierdo Arroyo (1992, p. 39) dividiu a sua proposta de estudo em: descrição indicativa, esquema-sumário e referências adotadas. Ele sumarizou os tópicos centrais dos estudos, identificando capítulos e seções detalhadamente, possibilitando sublinhar até as matrizes teóricas escolhidas.

Com base na Semiótica de Peirce, no eixo que trata da representação documental, o autor pontuou como primeiro trabalho necessário "Um marco semiótico para a teoria da representação". Em outros termos, apresentou uma exposição completa dos principais conceitos da Semiótica peirceana, expondo a divisão da Semiótica em Gramática Especulativa, Retórica Pura e Lógica Geral.(IZQUIERDO ARROYO, 1992, p. 39-40). Isso revela uma preocupação em articular seriamente os conceitos semióticos de Peirce com a Documentação, longe de propor uma suposta interdisciplinaridade, que apenas recebe e procurar aplicar o que seja possível.

A base da teoria da representação a ser estudada ancora-se na Semiótica de Peirce. Contudo, no segundo trabalho em curso, "Análise de conteúdo e representação documental", não indica a possibilidades de síntese com a Semiótica, preferindo pôr ênfase às contribuições do Gerativismo, Semântica Estrutural, Análise do Discurso e Linguística Textual.

O segundo eixo de trabalho trata dos estudos históricos agrupados na expressão Semiótica Documental. Izquierdo Arroyo (1992, p. 44-51) relacionou como discussões da história da Semiótica Documental a origem das seguintes problemáticas: a organização de textos, as operações de análise e síntese, a indexação, a descrição textual, a catalogação, a construção de repositórios bibliográficos, a classificação científica, a organização alfabética, a produção de glossários, dicionários, enciclopédias e tesauros, a hierarquização temática e o estabelecimento das relações associativas. Em resumo, o ensaio histórico da Semiótica Documental trataria de estabelecer um diálogo com os principais temas da análise documental de conteúdo.

---

[2] Recentemente, Lara (2014) examinou o teor desse artigo de Izquierdo Arroyo, de 1992, descreveu a sua estrutura e avaliou a sua importância para a Documentação, ressaltando que os trabalhos de Izquierdo Arroyo não foram muito divulgados, e por conseguinte, houve um prejuízo no conhecimento de sua obra.

Contudo, a formalização da Semiótica Documental apareceu pela primeira vez no projeto docente de Izquierdo Arroyo – documento a ser apresentado em concurso público seguindo o disposto no Real Decreto n. 1427 de 1986 - redigido para admissão como professor da Universidad de Murcia. Em entrevista concedida em 2014, Izquierdo Arroyo esclareceu que quando estava escrevendo os "Esquemas de Lingüística Documental", entre os anos de 1989 e 1990, deparou-se com o enfoque agregador da Semiótica: "Pero el hilo conductor de mi propuesta era ya la Semiótica peirceana. Para la elaboración del trabajo Sobre la transducción, había reunido en Burgos bastante material de Semiótica y Semiología." (IZQUIERDO ALONSO; IZQUIERDO ARROYO, 2014, p. 111).

Conquanto, é possível inferir que seu contato com o pensamento semiótico de Peirce mais dedicado foi no final de 1970.

> Instalado en la Filosofía del Lenguaje (de corte más bien analítico) y en la Semiótica – acababa de leer a Peirce -, consideré que la aproximación más adecuada sería la lógico-linguística (término este que ya empezaba a aparecer en algunas publicaciones recientes de entonces). A ello obedeció el titulo y orientación del libro Concepción lógico-linguística de la Documentación. (IZQUIERDO ALONSO; IZQUIERDO ARROYO, 2014, p. 111)

Em 1989, foi aprovado no concurso para professor em Documentação na Universidade de Múrcia. O concurso foi convocado através do ato administrativo da referida universidade, Resolução n. 1309, de 10 de janeiro de 1989, para formar parte do grupo de professores titulares para a Diplomatura em Biblioteconomia e Documentação. O concurso foi para a área (matéria) Análise e Linguística Documentais, contando com as disciplinas *Análisis documental*, *Lingüística documental I* (10 créditos) e *Lingüística documental II* (5 créditos). Foi este o contexto da elaboração do projeto docente, redigido em Málaga no ano de 1989 e que nasceu uma clara definição da Semiótica Documental.

As disciplinas *Lingüística documental I* e *Lingüística documental II*, segundo consta em seu livro "Esquemas de lingüística documental" (IZQUIERDO ARROYO, 1990), juntam-se a outras disciplinas em que se trabalhavam os temas semióticos, a saber: *Seminario de Estudios I*, *Seminario de Estudo II*, *Trabalho de Fin de Curso*. Esse arranjo de disci-

plinas congregou as reflexões originais sobre Semiótica e Linguística no campo da Documentação.

Por vezes, Izquierdo Arroyo oferecia também uma optativa chamada "Aspectos lógico-lingüísticos de la Documentación". Através dessas disciplinas foram expostos aos alunos as perspectivas da Semiótica Documental como uma disciplina que viria a subsumir a Linguística Documental. Nessa mesma universidade, Izquierdo Arroyo aposentou-se em 2012, após mais de 30 anos de magistério.

As bases teóricas de Izquierdo Arroyo recobrem diversos campos das humanidades, mas com uma base consistente em Filosofia e Letras.

> Del lado de la Documentación, me movieron alguns ideas redundantes en P. Otlet, J. Chaumier, M. Coyaud, Y. Courrier, M. Taube, J.C. Gardin, S.R. Ranganathan, D.J. Foskett, B.C. Vickery, D. Austin, De Grolier, Fugmann, I. Dahlberg y otros. Del lado de la formación en Filosofía, creo que puedo señalar tres referentes destacados, si bien menos conocidos en el ámbito documental: a) la "Grammatica Speculativa, sive De modis significandi", atribuida al Pseudo-Scoto/ Tomás de Erfut – entre otros tratados de semiótica medieval -; b) los escritos lógico-semióticos de Charles Sanders Peirce (el mejor conocedor de esos tratados medievales dentro del área anglosajona); y c) el pensamiento perspectivista de don José Ortega y Gasset. (IZQUIERDO ALONSO; IZQUIERDO ARROYO, 2014, p. 112)

Não se deve desmerecer as contribuições teóricas a sua formação intelectual dos pensadores Greimas, Todorov, Barthes, Kristeva, Derrida, Louis Hjelmslev, Lyons, Bühler, Jakobson e formalistas russos. Segundo Moreno Fernandez e Izquierdo Alonso (2014, p. 21), Izquierdo Arroyo pode ser considerado um filósofo e investigador que aporta contribuições no campo da Lógica formal, Filosofia da ciência e Filosofia da linguagem. Ele desenvolveu trabalhos em diversas áreas, Moreno Fernandez e Izquierdo Alonso (2014) abordaram o desenvolvimento do autor em três grandes eixos de atividades: os estudos ligados à teoria da Documentação; os estudos ligados à teoria da comunicação e Semiologia; os estudos relacionados à Semiótica documental.

> Tras su publicación em 1980 de su obra *Sobre la transducción, meditaciones semiológicas*, y desde sus primeras reflexiones a inicios de los 90 relacionadas con el *tratamiento documental de contenido* y

> el *tratamiento temático de la información*, formuló y conceptualizó, en el marco de un a semiótica estructural-funcional, su propuesta de *semiótica documental* y *definió sus modelos formales para la representación documental* y el estudio de los lenguajes documentales. (MORENO FERNANDEZ; IZQUIERDO ALONSO, 2014, p. 21, grifo dos autores)

Na Universidade de Múrcia, Izquierdo Arroyo, além das disciplinas já mencionadas que funcionaram como laboratório para os experimentos e sistematização da Semiótica Documental, vale a pena registrar o papel do grupo de pesquisa que leva este nome SemioDoc (Semiótica Documental). Fundado em 1991 junto à Universidad de Murcia, o grupo funcionou como "comunidad de práctica investigadora y docente[...]", que, segundo Moreno Fernandez e Izquierdo Alonso (2014, p. 30) abordou os seguintes temas: semiótica documental descritiva, modelos de tratamento documental do conteúdo, gestão de linguagens documentais, taxonomia, teoria da classificação, pragmática documental da representação e reconhecimento, gramáticas gerativas aplicadas à estruturação textual, sistema de processamento de discurso, modelos e técnicas de resumo, tratamento semântico de documentos icônicos e digitais etc. Somente por esta lista pode-se constatar a amplitude do espectro temático da Semiótica documental[3].

De acordo com Moreno Fernandez e Izquierdo Alonso (2014, p. 27) o livro de Izquierdo Arroyo, "Esquemas de lingüística documental" é um material obrigatório na discussão teórica da análise documental de conteúdo e da Semiótica Documental:

> Constituye una monografia de referencia obligatoria en los estudios sobre análisis documental de contenido y su herramienta específica: los lenguajes documentales. En ella se asientan las bases de la concepción científica de la LD, como disciplina teórico-práctica que se plantea el almacenamiento racional y ulterior recuperación del contenido analítico de cualesquiera documentos (definición pragmática).

[3] O grupo de pesquisa SEMIODOC continua ativo e contanto com a participação de professores, colaboradores e bolsistas. Mas informações sobre temas de pesquisa atuais podem ser levantadas em: <https://curie.um.es/curie/catalogo-ficha.du?seof_codigo=1&perf_codigo=10&cods=E053*02>.

A Semiótica Documental assenta-se marcadamente na concepção semiótica de Peirce e na divisão das dimensões da semiose desenvolvida por Morris. A estruturação da Semiótica Documental toma como pressuposto o problema lançado por García Gutiérrez, a saber: o problema da Linguística Documental, o qual deveria ser enfrentado nos primeiros anos da década de 1990 por todos os especialistas em Documentação que se ocupam da análise e da sistematização do conteúdo.

Examinando com mais acuidade, Izquierdo Arroyo (1990) apontou as características do objeto desta disciplina: "Ese objeto es algo de carácter semántico – el llamado "contenido", no muy felízmente – o semiótico, y por tal, distinto del plano material o "físico" o meramente "expresivo" (plano hjelmsleviano de la expresión)." (IZQUIERDO ARROYO, 1990, p. 42). Para o autor, a disciplina não se confunde com análise documental, pois considera esta expressão apenas voltada à descrição bibliográfica ou física dos documentos.

No que respeita a base teórica, como foi explicitado anteriormente, a Semiótica documental deriva da concepção semiótica, em primeiro lugar, de Peirce, e em segundo, de Charles Morris.

> Por tal entiendo algo más que un simples tratamiento "lingüístico", en el sentido próprio de este término en el uso y mención que de él hace la Teoría General del Lenguaje. El adjetivo "semiótico" (que provisionalmente puede leerse como "lógico-lingüístico") lo utilizo en la acepción promovida por Ch. S. PEIRCE; para quien la Semiótica (<u>Semeiotic</u>) aúna ternariamente las perspectivas de Gramática (aspecto lingüístico propiamente tal), la Lógica formal-material (aspecto lógico) y la Retórica, peculiarmente entendida por nuestro autor (aspecto comunicativo y pragmático). Con ello quiero también significar que – en consonancia con la teorización de Ch. MORRIS -, dicho tratamiento semiótico se mueve en tres dimensiones que bien se conocen: <u>sintáctica</u>, semántica y pragmática. Y subrayo la primera por cuanto entiendo que en si falta de consideración decidida radican los principales problemas de los lenguajes documentales más en uso. (IZQUIERDO ARROYO, 1990, p. 43)

A observação de Izquierdo Arroyo explica tanto as bases quanto o nível dos problemas em Documentação. Nesse caso, sobressai-se o grupo de problemas sintáticos. Diferente da Linguística documental que se

alinha à Linguística, apesar da notória influência linguística, a Semiótica Documental recorre à Semiótica geral de Peirce para projetar seus ramos e colocar-se os problemas de tratamento do conteúdo na Documentação sob este quadro de análise. Desenvolvendo um pouco mais a reflexão sobre a disciplina e apropriando-se das dimensões da semiose de Morris, Izquierdo Arroyo destacou alguns desenvolvimentos históricos da disciplina.

Izquierdo Arroyo (1990, p. 46) é tributário logicamente à García Gutiérrez pois este lançou à Documentação a questão da Linguística Documental, contudo propõe a abertura da acepção empregada à disciplina, resultado da reflexão e questionamento do rótulo até então utilizado. Depreendemos dos argumentos de Izquierdo Arroyo: 1) os documentos, de longe, não se limitam mais ao escrito-textual, dada a abundância e a variedade dos códigos e suportes documentais; 2) a questão semântica e pragmática mobiliza outros conhecimentos que não os tradicionalmente utilizados pela Documentação de recorte linguístico; 3) a existência de uma multiplicidade de códigos e de processos de tradução inter e entre códigos na Documentação. Portanto, já se está há algum tempo em um campo semiótico por excelência. Nesse sentido, não seria mais que adequado redefinir a disciplina Linguística Documental, e projetar uma disciplina mais robusta e coerente com as novas descobertas científicas e desenvolvimentos em Documentação.

Segundo Izquierdo Arroyo (1993, p. 200), a Semiótica Documental é o marco acolhedor das denominadas Ciências do Texto em sua aplicação ao tratamento documental. Por texto, Izquierdo Arroyo (1993, p. 201) definiu a representação física do discurso, escrito ou oral, e por Ciências do Texto compreende pelo menos a Linguística textual e as Ciências Cognitivas. Nesse caso, a Semiótica Documental já nasce com o reconhecimento das teorias cognitivas como linhas explicativa da leitura documental e do texto.

A definição de Semiótica tem como intenção principal reconstituir o papel da tríade recordada pelo autor de expressão/conteúdo/referente. Como observado, a proposição da Semiótica Documental ressalta a necessidade de uma leitura dos três planos, portanto, presumimos que se confira à teoria triádica do signo uma importância capital.

Além disso, assumindo a Semiótica peirceana como matriz, tal como se depreende da análise dos conceitos de Peirce utilizados por Izquierdo Arroyo (divisões da lógica, signo, segunda tricotomia, interpretante, realidade, mundo e coisa), o autor (IZQUIERDO ARROYO, 1993) afirma que a substituição se justifica porque a Semiótica cobre todos os tipos de signos, em que representam o discurso e podem ser símbolos, ícones e índices.

Em resumo, o autor concluiu o argumento sustentando que a Linguística não cobria a totalidade dos signos, essa tarefa é cumprida pela Semiótica (IZQUIERDO ARROYO, 1993, p. 202). Tal argumentação é coerente e se mostra eficaz dada a abrangência pretendida, por isso, adotou Semiótica no lugar de Linguística.

A disciplina em questão, "Linguística documental estendida", ou "Semiótica documental" pertenceria à teoria geral da Documentação (IZQUIERDO ARROYO, 1990, p. 49, p. 54). Talvez isso significasse que a Semiótica Documental, longe de ser um espaço de aplicações e experimentações semióticas, foi mais bem uma perspectiva particular da Documentação frente aos problemas do tratamento documental de conteúdo que atingem várias disciplinas e profissões. Assim, a Semiótica Documental poderia ser considerada um ramo da Documentação que ofereceria soluções aos problemas de tratamento de informação a outros campos, ao mesmo tempo em que se concentraria nos problemas gerais de processamento da linguagem na sociedade.

Contrastando a Semiótica Documental com a Linguística documental, notamos um descompasso dessa última com os problemas não linguísticos que foram potencializados com o uso de documentos imagéticos, audiovisuais e da hipermídia.

Essa aspecto também foi ressaltado por Moreno Fernandez e Izquierdo Alonso (2014, p. 27). A definição de tratamento documental entendido por Izquierdo Arroyo (1990, p. 55) compreende tanto os aspectos físicos quanto os de conteúdo do documento, não obstante, apenas este último seria objeto da Semiótica documental. É justamente dentro deste recorte epistemológico que encontramos o salto qualitativo da Semiótica

documental, o de considerar os processos de semiose em Documentação como o centro do debate.

> Sin embargo - aunque académicamente he apostado ya por la denominación "Lingüística Documental" - me inclino a proponer para nuestra disciplina una denominación más abarcadora de sus propósitos: SEMIÓTICA DOCUMENTAL. Así nombrada, aparecería como una especificación disciplinar de la Semiótica general, y habría de entenderse como ciencia que estudia la semiosis documental. El la concepción de PEIRCE y MORRIS, entre otros, la semiosis es el proceso de comunicación general de que se ocupa la Semiótica, y en él los signos se consideran em sus tres vertientes: sintáctica, semántica y pragmática. (IZQUIERDO ARROYO, 1990, p. 63-64)

Infelizmente esta obra não foi divulgada na época de sua publicação na literatura científica da Documentação nos países sul-americanos, e nos estudos da Documentação de língua portuguesa, em especial, no contexto brasileiro. A substituição terminológica da Linguística Documental daria lugar, por uma questão lógica, à Semiótica Documental. Porém, o que sucedeu foi o uso da expressão Linguística Documental pela comunidade de Documentação, especialmente no Brasil, quando, em realidade, as condições objetivas (documentos em vários códigos) e os temas de investigação já exigiam uma reformulação, pois já se aproximavam dos problemas semióticos.

A Semiótica Documental, como exposto, seria o marco acolhedor das preocupações do tratamento documental del conteúdo, abarcaria conjuntamente Linguística, Ciências Cognitivas e Ciências da Comunicação, incluídas como ramos da Semiótica geral.

O trabalho de Izquierdo Arroyo em Semiótica Documental, a despeito de o nível avançado de sistematização, ainda deixou pontos a avançar e exigirão dos especialistas em Documentação propostas coerentes para seu desenvolvimento. Izquierdo Arroyo tem chamado a atenção da Documentação para o problema da variedade de códigos que se apresentava já final dos anos 1980 e para o qual não se tem ainda um campo coerente ou abordagem unificada na área. Atualmente, ainda se atua como se o problema do conteúdo fosse apenas uma questão de tratamento textual, e o que não possui essa característica simbólica deve adaptar-se, quando, em

realidade, os diversos códigos icônicos e indiciais já reclamavam um espaço próprio no consumo de informação na sociedade.

Izquierdo Alonso e Izquierdo Arroyo (2014, p. 116) argumentam que as coleções de orais e os documentos icônicos, as imagens, não têm tratamento adequados. Esse é um dos desafios atuais da Documentação, em especial, da Semiótica documental, pois os documentos imagéticos e sonoros exigem competências que estão além da teoria linguística. Aqui há uma clara separação entre o que pode produzir a Linguística Documental e o que preocupa a Semiótica documental de Izquierdo Arroyo.

A Semiótica Documental demonstrou cabalmente que os problemas da Linguística Documental deveriam ser bem outros, mais condicionados à realidade da informação e aos códigos utilizados. Os problemas elencados pela Linguística documental, acreditamos, já surgiram defasados no tempo e parece que apenas Izquierdo Arroyo notou esse fenômeno. Com a análise podemos avançar na compreensão das ideias de Izquierdo Arroyo, e começar a responder a precisa constatação de Lara (2014): "No Brasil, o acesso aos textos de Izquierdo Arroyo é bastante incompleto, e essa é uma das razões pela qual sua produção é pouco conhecida."

Assim, a Semiótica Documental se configura como um campo híbrido que recebe contribuições filosóficas e científicas destinadas a responder aos problemas teóricos e aplicados do tratamento de documentos em linguagem verbal e não verbal. Sem uma postura semiótica não poderíamos incluir o percurso da leitura para recuperação da informação, tal como fomentado pelos estudos de Agustín Lacruz (2006, 2015), Manini (2001, 2004) e Pato (2014). Em síntese, devemos rediscutir o que a noção ampliada de leitura dentro do escopo da indexação implica, em especial, a sua aplicação a códigos não verbais.

4 Considerações finais

A leitura documental tem recebido muitas contribuições provenientes de uma diversidade de áreas de pesquisa, nesse sentido, uma suposta aglutinação de conceitos linguísticos não seria suficiente para explicar este processo tão complexo.

Por esta razão, a Semiótica Documental pode ser a mais adequada estrutura de confluência conceitual a ser desenvolvida no interior da Documentação no sentido de propor como um de seus objetos básicos a leitura com fins documentais. As vertentes da Linguística, das Ciências Cognitivas, das Ciências da Comunicação, entre outras, seriam exploradas, no arranjo da Semiótica Documental. Os estudos da leitura documental, entre tantos outros com tratamento pulverizado na Documentação, pertenceriam aos interesses da Semiótica Documental.

A leitura documental contaria com uma riqueza científica derivada de estudos aplicados e princípios teóricos para compreender a semiose que ocorre em sistemas de informação, desde a representação do conteúdo do documento até a apropriação da informação pelos usuários.

Referências

ABREU, J. G.; MONTEIRO, S. D. Matrizes da linguagem e a organização virtual do conhecimento. *Ciência da Informação*, Brasília, v. 29, n. 2, p. 9-26, maio/ago. 2010.

AGUSTÍN LACRUZ, M. C. *Análisis documental de contenido del retrato pictórico*: propuesta epistemológica y metodológica aplicada a la obra de Francisco de Goya. Cartagena: Ayuntamiento, Concejaría De Cultura, 2006.

AGUSTÍN LACRUZ, M. C. Lectura de las imágenes fotográficas orientada hacia la representación documental. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, Florianópolis, v. 20, n. esp. 1, p. 55-88, Fev., 2015. Disponível em:<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2015v20nesp1p55/28639>. Acesso em: 27/01/2015.

ALMEIDA, C. C. Desafios para uma teoria peirceana da organização da informação e do conhecimento. In: SILVA, F. C. C.; SALES, R.(Org.). *Cenários da organização do conhecimento*: linguagens documentárias em cena. Brasília: Thesaurus, 2011, p. 71-86.

ALMEIDA, C. C.; GARCÍA MARCO, F.J. Aportaciones Semióticas de la Documentación en España a la Organización Del Conocimiento: un análisis preliminar. In: *XII Congreso ISKO España y II Congreso ISKO España-Portugal.* Murcia: Universidad de Murcia, 2015.

CHAUMIER, J. *As técnicas documentais*. Lisboa: Publicações Europa-América, 1971.

CUNHA, I. M. R. F. (Coord.). *Análise documentária*: considerações teóricas e experimentações. São Paulo: FEBAB, 1989.

FRIEDMAN, A.; THELLEFSEN, M. Concept theory and semiotics in knowledge organization. *Journal of Documentation*, London, v. 67, n. 4, p. 644-674, 2011.

FUJITA, M. S. L. A leitura documentária na perspectiva de suas variáveis: leitor-texto-contexto. *DataGramaZero*, v. 5, n. 4, p. 00, 2004. Disponível em: <http://basessibi.c3sl.ufpr.br/brapci/v/a/7547>. Acesso em: 27 Maio 2017.

GARCÍA GUTIÉRREZ, A. L. *Estructura lingüística de la documentación*. Murcia: Universidad de Murcia, 1990.

GARCÍA GUTIÉRREZ, A. L. *Lingüística documental*: aplicación a la documentación de la comunicación social. Barcelona: Mitre, DL, 1984.

GARCÍA MARCO, F. J. Reflexiones entorno a las aportaciones de José María Izquierdo Arroyo a la representación y la organización del conocimiento: ambición teórica, perspectiva humanística y compromiso académico. *Scire*, Zaragoza, v. 20, n. 1, p.13-19, ene. /jun. 2014.

GARDIN, J. C. *Les analyses de discours*. Neuchâtel: Delachaux et Niestlê, 1974.

GARDIN, J. C. Éléments d´um modele pour la description de lexiques documentaires. *Bulletin des Bibliothèques de France*, v. 11, n. 5, p. 171-182, 1966.

GARDIN, J. C. Vers une épistémologie pratique en sciences humaines. In: GARDIN, J. C. et. al. *La logique du plausible*. Paris: La Maison des Sciences de l'Homme, 1987. p. 27-102.

HJØRLAND, B. Fundamentals of knowledge organization. *Knowledge Organization*, v. 30, n. 2, p. 87-111, 2003.

IZQUIERDO ALONSO, M.; IZQUIERDO ARROYO, J. M. Entrevista a José María Izquierdo Arroyo, realizada por Mónica Izquierdo Alonso em Alcalá de Henares (Madrid), junio de 2014. *Scire*, Zaragoza, v. 20, n. 1, p.109-116, ene. /jun. 2014.

IZQUIERDO ARROYO, J. M. *Esquemas de lingüística documental.* Barcelona: Promociones y Publicaciones Universitarias, 1990. Tomo II - p. 243-506.

IZQUIERDO ARROYO, J. M. *Esquemas de lingüística documental*. Barcelona: Promociones y Publicaciones Universitarias, 1990. Tomo III - p. 507-871.

IZQUIERDO ARROYO, J. M. *Esquemas de lingüística documental.* Barcelona: Promociones y Publicaciones Universitarias, 1990.Tomo I - p. I-242.

IZQUIERDO ARROYO, J. M. La ciencia de la búsqueda documental secundaría. *Documentación de las Ciencias de la Información*, Madrid, n. 13, p. 87-111, 1990.

IZQUIERDO ARROYO, J. M. Cuatro trabajos en curso. *Documentación de las Ciencias de la Información*, Madrid, n. 15, p. 35-65, 1992.

IZQUIERDO ARROYO, J. M. De la semiótica del discurso a la semiótica documental. In: MORENO GONZÁLEZ, J. A. *Aplicación de las ciencias del texto al resumen documental*. Madrid: Universidad Carlos III, 1993. p. 199-216.

IZQUIERDO ARROYO, J. M. *La organización documental del conocimiento*. Madrid: Tecnidoc, 1995.

LANCASTER, F. W. *Indexação e resumos*. Brasília: Briquet de Lemos/Livros, 1993.

LARA, M. L. G. Linguagem documentária e terminologia. *Transinformação*, Campinas, v. 16, n. 3, p. 231-240, set./dez. 2004.

LARA, M. L. G. ; TÁLAMO, M. F. G. M. Lingüística documentária e terminologia: experiência didática na interface das disciplinas. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 7., 2006, Marília. *Anais...* Marília: FFC/UNESP-PUBLICAÇÕES, 2006. 1 CD-ROM.

LARA, M. L. G. ; TÁLAMO, M. F. G. M. Uma experiência na interface Linguística Documentária e Terminologia. *DataGramaZero*: Revista de Ciência da Informação, Rio de Janeiro, v. 8, n. 5, out. 2007. Disponível em: <http://www.dgz.org.br/out07/Art_01.htm>. Acesso em: 03 abr. 2013.

LARA, M. L. G. Algumas contribuições da semiologia e da semiótica para a análise das linguagens documentárias. *Ciência da Informação*, Brasília, v. 22, n. 3, p. 223-226, set./dez. 1993.

LARA, M. L. G. Aplicação de um modelo de análise documentária à literatura sócio-econômica. In: CUNHA, I. M. R. F. (Coord.). *Análise documentária*: considerações teóricas e experimentações. São Paulo: FEBAB, 1989. p. 131-182.

LARA, M. L. G. Conceitos lingüísticos fundamentais para a organização e disseminação de informações. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5., 2003, Belo Horizonte. *Anais...* Belo Horizonte: UFMG, 2003. 1 CD-ROM.

LARA, M. L. G. É possível falar em signo e semiose documentária? *Encontros Bibli*: Revista Eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, 2º n. especial, p. 18-29, 2º sem. 2006. Disponível em: <http://www.periodicos.ufsc.br/index.php/eb>. Acesso em: 15 out. 2006.

LARA, M. L. G. O unicórnio (o rinoceronte, o ornitorrinco...), a análise documentária e a linguagem documentária. *DataGramaZero* – Revista de Ciência da Informação, Rio de Janeiro, v. 2, n. 6, dez. 2001. Disponível em: <http://www.dgz.org.br/dez01/Art_03.htm>. Acesso em: 25 ago. 2007.

LARA, M. L. G. Propostas de tipologias de KOS: uma análise das referências de formas dominantes de organização do conhecimento. *Encontros Bibli*: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, v. 20, n. esp. 1, p. 89-107, Fev., 2015. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/15182924.2015v20nesp1p89/28637>. Acesso em: 27/01/2015.

LARA, M. L. G. *Representação e linguagens documentárias*: bases teórico-metodológicas. 1999. 207 f. São Paulo. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes. Universidade de São Paulo.

LARA, M. L. G. Sobre “Cuatro trabajos en curso” de José María Izquierdo Arroyo. *Scire*, Zaragoza, v. 20, n. 1, p. 91-98, ene./jun. 2014.

LARA, M. L. G. Uma teoria terminológica para um conceito contemporâneo de informação documentária. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 6., 2005, Florianópolis. *Anais...* Florianópolis: UFSC, 2005. 1 CD-ROM.

MAI, J-E. Semiotics and indexing: an analysis of the subject indexing process. *Journal of Documentation,* London, v. 57, n. 5, p. 591-522, sep. 2001.

MAI, J-E. The concept of subject in a semiotic light. In: SCHWARTS, C.; RORVIG, M. (Ed.). *Digital collections*: implications for users, funders, developers and maintainers. Medford, NJ: Information Today, 1997a. p. 54-64. (Proceedins of the ASIS Annual Meeting; 34).

MAI, J-E. The concept of subject: on problems in indexing. In: McILWAINE, I. C. (Ed.). *Knowledge organization for information retrieval*: 6th International Study Conference on Classification Research. The Hague: FID, 1997b. p. 60-67. (FID, n. 716).

MAI, J-E. *The subject indexing process*: an investigation of problems in knowledge representation. 2000. 344 f. Dissertation (Doctor of Philosophy) – Faculty of Graduate School of Library and Information Science. The University of Texas at Austin.

MANINI, M. P. Análise documentária de fotografias: leitura de imagens incluindo sua dimensão expressiva. *Cenário Arquivístico*: Revista da Associação Brasiliense de Arquivologia, Brasília, v. 3, n. 1, p. 16-28, 2004.

MANINI, M. P. Análise documentária de imagens. *Informação e Sociedade*: Estudos, João Pessoa, v. 11, n. 1, 2001. Disponível em: <http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/313/236>. Acesso em: 22 set. 2007.

MORENO FERNÁNDEZ, L.M.; IZQUIERDO ALONSO, M. El pensamiento de José María Izquierdo Arroyo em la organización y representación del conocimiento: una sistematización desde fundamentos filosóficos y semióticos. *Scire*, Zaragoza, v. 20, n. 1, p.21-32, ene. /jun. 2014.

MOURA, M. A. ; SILVA, A. P.; AMORIM, V. R. A concepção e o uso das linguagens de indexação face às contribuições da Semiótica e da Semiologia. *Informação e Sociedade*: Estudos, João Pessoa, PB, v. 12, n. 1, p. 1-22, 2002. Disponível em: <http://www.informacaoesociedade.ufpb.br/issuev12n102.htm>. Acesso em: 13 out. 2005.

MOURA, M. A. Ciência da informação e semiótica: conexão de saberes. *Encontros Bibli*: Revista Eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, 2º n. especial, p. 1-17, 2º sem. 2006. Disponível em: <http://www.periodicos.ufsc.br/index.php/eb>. Acesso em: 15 out. 2006.

MOURA, M. A. Interoperabilidade semântica e ontologia semiótica: a construção e o compartilhamento de conceitos científicos em ambientes colaborativos online. *Informação e Informação*, Londrina, v. 16, n. Esp., p. 165-179, jan./jun. 2011.

MOURA, M. A. Signi-fica ou signi-vai? as teorias da significação no campo da Ciência da Informação. In: REIS, A. S. ; CABRAL, A. M. (Org.) *Informação, cultura e sociedade*: interlocuções e perspectivas. Belo Horizonte: Novatus, 2007. p. 61-80.

PATO, P. R. G. Ícone, índice e símbolo, fundamentos para ler e organizar a informação em imagens. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, XV., 2014, Belo Horizonte, MG. *Anais...* Belo Horizonte, MG: UFMG, 2014. p. 488-508.

RABER, D.; BUDD, J. M. Information as sign: semiotics and information science. *Journal of Documentation*, London, v. 59, n.5, p. 507-522, 2003.

REAL DECRETO 1497/1987, de 27 de noviembre, por el que se establecen directrices generales comunes de los planes de estudio de los títulos universitarios de carácter oficial y validez en todo el territorio nacional. Disponível em: <https://www.boe.es/diario_boe/txt.php?id=BOE-A-1987-27707>. Acesso em: 21/10/2014.

THELLEFSEN, T. L. Pragmaticism and the role of terminology. *Impact*: an electronic journal on formalisation in text, media and language, abr. 2003. Disponível em: <http://www.impact.hum.auc.dk>. Acesso em: 05 maio 2007.

THELLEFSEN, T. L. ; THELLEFSEN, M. M. Pragmatic semiotics and knowledge organization. *Knowledge Organization*, v. 31, n. 3, p. 177-187, 2004.

THELLEFSEN, T. L. *Fundamental signs and significance-effects*: a semeiotic outline of fundamental signs, significance-effects, knowledge profiling and their use in knowledge organization and branding. 2009. 284 f. Doctoral Dissertation – Departament of internationale Culture and Comunication Studies. Copenhagen Business School.

THELLEFSEN, T. L. Knowledge profiling: the basis for knowledge organization. *Library Trends*, v. 52, n. 3, p. 507-514, winter 2004.

THELLEFSEN, T. L. Semiotic knowledge organization: theory and method development. *Semiotica*, v. 142, n. 1 / 4, p. 71-90, 2002.

Análise semiótica da leitura documental no processo de indexação de assunto

Luciana Beatriz Piovezan

1 Introdução

A leitura em Biblioteconomia possui papel importante na atuação profissional, dado que o tratamento que recebem os documentos, com vistas à disponibilização para acesso e uso, necessita da leitura para ser realizado. Desse modo, a leitura insere-se em diversas atividades realizadas em unidades de informação, tais como, a catalogação, a elaboração de resumos, a classificação e a indexação. A leitura é fundamental para as atividades de tratamento temático da informação documental, em especial para a indexação, por causa da necessidade de representação do conteúdo temático dos documentos.

Em Unidades de Informação existem diversas tipologias documentais – livros, revistas, filmes, álbuns musicais, entre tantos outros – em diversos suportes físicos, todos requerendo do profissional tratamento adequado. Aqui, entretanto, abordaremos as questões referentes à leitura para tratamento de documentos textuais, que apresentam linguagem escrita. Estas reflexões referem-se principalmente a livros e artigos, ambos de caráter científico.

As interseções entre a semiótica peirceana e a organização da informação, especialmente no que tange à indexação, têm sido abordadas na literatura recente de Ciência da Informação, como por exemplo, em

Souza e Almeida (2012) que abordou a relação da semiótica e da indexação, ressaltando as questões da representação e do referente; em Redígolo e Almeida (2012) onde se encontra um estudo da análise de assunto por meio da semiótica de Peirce abordando os conceitos de hábito, experiência, representação e destacando também os tipos de signos que a análise de assunto faz surgir. Do mesmo modo, encontramos em Dal'Evedove, Almeida e Fujita (2013) uma análise do processo de indexação que o vislumbra como um processo semioticamente estruturado, dado que se faz por meio da interpretação; ainda em Almeida; Fujita e Reis (2013) os autores procurar tratar a relação da indexação com o segundo ramo da Semiótica, a Lógica Pura, argumentando que a indexação é um processo inferencial por considerar a interação entre tipos diversos de signo neste processo com intenção de representação da informação.

Percebemos que estes trabalhos anteriores tangenciam a questão da leitura documental, apresentando-a ora implícita (cf. SOUZA; ALMEIDA, 2012) ora explicitamente (cf. REDÍGOLO; ALMEIDA, 2012; ALMEIDA; FUJITA; REIS, 2013), porém sem efetivamente oferecer subsídios para sua compreensão. Suscitados, também, por essa vacância literária, neste trabalho pretendemos abordar a possibilidade de compreensão da leitura documental meio da perspectiva semiótica de Peirce. Para traçar este caminho, iniciaremos expondo os principais aspectos da leitura documental, apresentaremos a semiótica peirceana, abordando as principais contribuições destes estudos para a compreensão das questões informacionais e, em seguida, serão expostas as contribuições identificadas no estudo da Semiótica para a compreensão da leitura documental.

2 Aspectos relativos à leitura

Passaremos a apresentação das questões relativas à leitura documental, iniciando pela exposição do processo de leitura, como percebido pelos modelos ascendente, descendente e interativo, de modo que poderemos passar ao entendimento da leitura documental ou leitura para a realização dos procedimentos de representação de assunto, compreendida como integrante do processo de indexação.

As primeiras concepções sobre a leitura a descrevem como a decodificação de um código, a escrita, sendo necessário ao leitor apenas a habilidade técnica de identificar os caracteres e sua relação linear entre letras, palavras, frases e textos, não fazendo menção à relação entre o leitor com o texto para adiante disso. Esta noção de leitura parece estar relacionada à perspectiva semiológica no que diz respeito à visão estruturalista da linguagem.

Essa perspectiva linear sobre a leitura é compreendida dentro do modelo ascendente de leitura – *bottom up* ou modelo serial de Gough (FUJITA; NARDI; FAGUNDES, 2003). Neste modelo centrado no texto, se considera que o leitor realiza um processo ascendente e sequencial de identificação e retirada do significado do texto para sua apreensão, partindo das partes para seu todo e realizando processos indutivos em que recebe o significado do texto.

Em contrapartida, o modelo descendente de leitura – modelo *top down*, oferece uma perspectiva na qual o leitor é o centro do modelo e, ao ler, realiza processos dedutivos partindo das unidades significativas do texto e relacionando-as com seus conhecimentos prévios para elaborar hipóteses e verificá-las no texto. A compreensão do texto, neste caso, é permitida pela introdução de conhecimento extralinguístico, de modo que o leitor não necessita de todos dados presentes no texto.

No presente momento, os pressupostos acerca da leitura englobam questões mais complexas e humanísticas, ressaltando o envolvimento cognitivo do leitor. A leitura, assim, é aceita como um processo de interação entre o conhecimento prévio do leitor com o conteúdo do texto para construção dos sentidos do texto (NEVES; RAMALHO; FIGUEIREDO, 2008). Essa perspectiva voltada à compreensão da leitura implica em que o significado do texto é construído pelo leitor de modo relacional, pois o texto escrito oferece diferentes possibilidades interpretativas.

Essa concepção de leitura é aceita no modelo denominado interativo, no qual se entende que os processos descritos pelos modelos ascendente e descendente são mais do que inversos, são complementares (FUJITA, 2004). A leitura, assim entendida, compreende procedimentos coordenados que incluem operações cognitivas, linguísticas e de percepção

visual que induzirão à compreensão do texto e da reconstrução de seu conteúdo pelo leitor.

Essa perspectiva interacionista compreende, portanto que "o sentido não precede o texto, não está nele depositado, nem é uma proposição pronta, acabada. É no texto sim, mas através do ato de leitura que se produzem os sentidos" (DUMONT, 2006, p. 12). De modo que para esta compreensão, o leitor deve ativar modelos mentais apreendidos em sua vivência pessoal para tornar a leitura um processo ativo.

A presença dos esquemas – "[...] estruturas cognitivas que dizem respeito ao conjunto de conhecimentos armazenados em sequência temporal ou causal, nos quais são mantidos os conjuntos de características dos objetos e seres que nos rodeiam" (NEVES, 2007, p. 3) – do leitor permite que ele vá criando hipóteses com base em inferências sobre o texto e, caso as confirme, estas são mantidas e mais desenvolvidas e, caso não possa confirmar suas hipóteses, o leitor as descarta e constrói novas hipóteses com base nos dados apreendidos do texto, alternando entre os processos ascendente e descendente.

A leitura é entendida, desse modo, como um ato interativo, dependente de suas variáveis, sendo elas, leitor e texto, assim como do contexto de leitura. Tais variáveis estão sujeitas a diversas condições que influenciam o processo de leitura (FUJITA, 2004). Nesse sentido, Neves (2007, p. 2) define a leitura de um texto como uma atividade cognitiva que

> [...] é também um processo interativo, em que o leitor e o autor se relacionam mutuamente, sendo o texto o agente desta vinculação. O autor, quando produz o texto, estabelece uma coerência textual que é recuperada no momento da leitura. A reconstituição da coerência de um texto depende de processos cognitivos construídos pelo leitor/usuário para sua compreensão.

Entendemos que a leitura é, portanto, uma prática social, relacionada a aspectos culturais envolvidos no cotidiano da sociedade, permitindo o compartilhamento de informação. Da mesma maneira a leitura é uma atividade subjetiva, que envolve aspectos cognitivos e emocionais relativos ao indivíduo leitor para construção do sentido de um texto. Assim como, a leitura é também, a principal forma de aces-

so ao conteúdo documental majoritariamente composto de documentos bibliográficos.

A leitura documental realizada pelo bibliotecário é a fase inicial do processo de tratamento dos documentos para sua inclusão em um sistema de informação e tem por finalidade a identificação de conceitos representativos do conteúdo temático do documento para sua representação. Moura (2006) lembra que este procedimento ativa outros dois, a análise de assunto e a indexação, e busca recompor a informação potencial do documento.

De acordo com Moura (2006) as práticas de leitura do bibliotecário podem levar a dois tipos de apropriação textual, a apropriação terminológica e a apropriação conceitual. A primeira se refere a uma assimilação por parte do bibliotecário dos termos recorrentes na sua área de atuação, porém sem compreender as relações teóricas, conceituais e metodológicas existentes. Já a segunda se refere à assimilação consciente dos conceitos e das relações conceituais existentes na área de conhecimento, realizando as atividades de organização do conhecimento de forma mais significativa e mais benéfica ao sistema e seus usuários.

Para que a leitura documental propicie a apropriação conceitual, consciente e significativa é necessário ao bibliotecário que tenha "[...] condições específicas como conhecimento prévio e de mundo, instrumentos, estratégias, e conhecimento de estruturas e tipos textuais e suas características" (REDÍGOLO; FUJITA, 2015, p. 357).

Dessa forma propõe-se que o bibliotecário seja visto como um elemento a ser considerado no estudo de metodologias de indexação, sendo o objetivo principal da formação do profissional atuante no tratamento temático "formá-lo ou capacitá-lo para uma leitura profissional, onde seus objetivos profissionais sobrepõem-se aos objetivos pessoais de leitor" (REDÍGOLO, 2007, p. 5).

Isto porque é impossível para profissionais de informação realizar a leitura do texto completo de todos os documentos que se apresentam para sua análise, por questões de operacionalização do trabalho. Desta forma, para uma melhor exploração textual, é necessário ao leitor profissional o conhecimento prévio de superestruturas ou esquemas textuais, as estru-

turas textuais de diferentes tipologias documentais, e a adoção de uma metodologia de abordagem ao documento, de maneira a permitir o alcance dos objetivos de leitura documental.

Em ambientes informacionais, o processo de indexação acontece em dois momentos, durante a inserção de um documento no sistema de informação e quando o sistema recebe um pedido de informação, em ambas as situações, tem por objetivo identificar a necessidade informacional do usuário. E não é excessivo ressaltar que a importância da leitura documental é devida a que a representação errônea ou incompleta dos documentos no processo de indexação irá refletir na qualidade da recuperação da informação e, portanto, na qualidade dos sistemas de recuperação de informação (MAI, 2001).

Gil Leiva (1999) observou a presença na literatura de biblioteconomia e ciência da informação de distintas definições para a indexação. Em geral tais definições procuram com maior ou menor completude caracterizar a indexação como um processo de análise, condensação e representação de conteúdos documentais. Tais operações são entendidas como etapas do processo de indexação.

Compreendendo, então, a indexação como um processo que possui três etapas – a determinação do conteúdo temático de um item, a análise conceitual para decidir quais aspectos de um item devem ser representados no registro bibliográfico e a tradução dos conceitos ou aspectos em um vocabulário controlado (OLSON; BOLL, 2001, p. 87), percebe-se que a leitura documental perpassa e possibilita esse processo, viabilizando a análise de assunto.

Deste modo, de acordo com Fujita (2013) existem dois momentos principais na leitura documental, sendo estes, a *identificação de conceitos* e a *seleção de conceitos* representativos do documento analisado. No primeiro momento, a identificação de conceitos, o bibliotecário realiza um exame lógico do texto de modo a selecionar conceitos representativos de seu conteúdo (FUJITA, 2003). Nesta fase, o bibliotecário realizará a determinação da tematicidade intrínseca do documento, a determinação de seu assunto mais significativo atendendo a composição do autor (FUJITA, 2013).

O segundo momento durante a leitura documental é a *seleção de conceitos*, onde o profissional considera os objetivos de leitura documental, seus interesses profissionais de leitura, realizando a determinação da tematicidade do documento, a representação do assunto mais significativo mais relevante ao contexto do usuário para o qual o documento se destina (FUJITA, 2013). Desta forma, é possível que conceitos identificados na primeira etapa não sejam selecionados para representar o conteúdo temático do documento.

Fujita (2013, p. 47) considera assim, que a análise de assunto consiste de quatro fases "a) determinação da tematicidade intrínseca, b) identificação de conceitos, c) seleção de conceitos e, d) determinação da tematicidade extrínseca".

Para a realização da leitura documental o bibliotecário irá ativar seu conhecimento prévio, composto por entidades tanto linguísticas quanto conceituais e que pode ser linguístico, textual, conhecimento de mundo e profissional, e suas estratégias de leitura.

Tais estratégias se dividem em dois grupos, as cognitivas, que se tratam de comportamentos inconscientes em uma leitura fluida e as metacognitivas, comportamentos conscientes do leitor em relação ao modo com que realiza a leitura. Quanto menos estratégias metacognitivas sejam exigidas pelo texto, mais fácil será a leitura, em contrapartida, a leitura metacognitiva conduz a maior compreensão do texto (NEVES, 2011).

Redígolo e Fujita (2015, p. 365) destacam ainda que "as estratégias metacognitivas demonstram que o leitor tem consciência dos objetivos da leitura e utiliza-se de alguns artifícios para realiza-la da melhor maneira, tentando superar suas dificuldades e explorando seus conhecimentos".

É em razão do objetivo já citado da indexação, identificar a necessidade de informação do usuário de modo a permitir a recuperação por assunto, que a indexação deve ser uma atividade personalizada, como afirma Lancaster (2004), para que a indexação ajuste-se com precisão aos interesses do grupo de usuários. O mesmo autor lembra a concordância de Hjørland (2001) quanto a essa questão, ao afirmar que "Uma vez que qualquer documento pode, em princípio, proporcionar respostas a uma infinidade de questões, as análises de assuntos devem estabelecer priorida-

des baseadas nos grupos de usuários específicos atendidos" (HJØRLAND, 2001, p. 776).

Hjørland (2001) definiu o assunto de um documento como as potencialidades epistemológicas (ou informativas) do documento. Adotaremos, apoiados no autor mencionado, esta noção de assunto enquanto potencial informativo de um documento.

Buscamos, portanto, a compreensão da filosofia Peirceana para refletir sobre a leitura profissional em indexação para representação do conteúdo temático de documentos.

3 A semiótica de Peirce

De modo sucinto, a semiótica deriva do exercício filosófico de Charles Sanders Peirce para compreender a realidade. Por meio da abstração sobre significado e linguagem, Peirce propõe o estudo do signo enquanto elemento triádico, composto por representamen, objeto e interpretante, em oposição à compreensão binária proposta por Ferdinand de Saussure. Tendo como fundamento a ideia de que a interpretação, que ele denomina semiose, é que cria o signo, o autor defende que qualquer coisa pode ser um signo, contanto que cause um efeito interpretativo em uma mente científica. Desta forma, a teoria de Peirce não se atém a linguagem verbal, como faz a teoria saussuriana, mas trata de como os significados são atribuídos aos signos em um sentido amplo.

Um dos conceitos centrais no quadro conceitual proposto por Peirce, semiose vem a ser o processo de construir significados por meio de signos:

> É importante que se entenda o que quero significar por semiose [...] uma ação, ou influência, que é, ou envolve, uma cooperação de três sujeitos, tais como um signo, seu objeto, e seu interpretante, essa tríplice relativa influência não sendo de modo algum resolúvel em ações entre pares. {Sémeiösis} em grego do período romano, desde o tempo de Cícero, se relembro bem, significa a ação de qualquer espécie de signo; e minha definição confere a qualquer coisa que assim atue o título de um "signo" (PEIRCE, 1980 p. 133-134).

Podemos, então, afirmar que o signo está para algo que interpreto como sendo análogo à minha memória daquele algo.

No processo de interpretação, Peirce distingue, principalmente, três tipos de interpretantes que são produzidos pela semiose, o interpretante imediato, o interpretante dinâmico e o interpretante final. O Interpretante Imediato pode ser compreendido como o interpretante representado no signo, assim o Interpretante Dinâmico seria o efeito de fato produzido na mente pela ação do signo e o Interpretante Final seria o objetivo de toda semiose, o propósito a ser alcançado por meio do desenvolvimento do pensamento (PEIRCE, 2005, p. 168-169).

Conforme expõe Lara, o interpretante "[...] também depende necessariamente do domínio, do discurso, das condições de enunciação e da experiência colateral do receptor" (LARA, 2004, p. 235-236).

Quanto ao objeto do signo, Peirce apresenta-os em dois tipos, Objeto Imediato, o objeto do modo como representado no signo, e Objeto Dinâmico, o referente em si, elemento indicado pelo signo. Em sua tentativa de fazer compreender tal distinção, o autor lembra que "O signo pode apenas representar o Objeto e referir-se a ele. Não pode proporcionar familiaridade ou reconhecimento desse Objeto [...] ele pressupõe uma familiaridade com algo a fim de veicular alguma informação ulterior sobre esse algo" (PEIRCE, 2010, p. 47). Para compreender o que é o Objeto Dinâmico não basta a experiência mediada pelo signo, faz-se necessária a existência da experiência colateral, ou seja, conhecer o objeto em si, para além do signo que o representa.

Para Lara (1993, p. 225) também a construção do significado está relacionada à experiência colateral, de modo que, "se essa experiência ensinou ao seu intérprete o que é determinada coisa, ou o que determinada palavra significa em uma dada sentença, sua interpretação será diferenciada". O objeto determina o signo, de maneira geral, e o interpretante de forma mais específica, e sendo assim, a experiência tida com o objeto é também determinante para a construção do interpretante. Como afirma Peirce (2010, p. 168, *destaques do autor*) "[...] o Signo *não pode* exprimir, ele pode apenas *indicar*, deixando ao intérprete a tarefa de descobri-lo por *experiência colateral*".

Para Peirce a função mais perceptível do pensamento é atingir um estado que ele denomina de crença, esse estado é a etapa final de um processo sistêmico que advém de uma perturbação que causa um estímulo - a dúvida, ou seja, a ação do pensamento inicia-se com o confronto com a dúvida e cessa ao chegar à crença.

Convém estabelecer que em Peirce (1975, p. 77) a crença "é a indicação mais ou menos segura de se ter estabelecido em nossa natureza uma tendência que determinará nossas ações. A dúvida nunca se acompanha de tal efeito". A crença possui três características "Primeiro é algo de que estamos cientes; segundo, aplaca a irritação da dúvida; e, terceiro, envolve o surgimento, em nossa natureza, de uma regra de ação, ou, digamos com brevidade, o surgimento de um *hábito*" (PEIRCE, 1975, p. 56).

Já a função final do pensamento seria a de construir hábitos de agir, os hábitos tem como função produzirem resultados sensíveis, portanto os hábitos seriam um modo de agir em uma dada circunstância para alcançar um determinado resultado desejado. Este modo de agir não deve ser definitivo, mas será sempre questionado pelo surgimento de dúvidas.

Para Peirce (2010, p. 48) toda esta construção está relacionada à existência de conhecimento prévio, pois em sua teoria, o autor deixa claro que a apreensão dos significados não é possível sem a determinação de sua relação com o Objeto, ou seja, se o indivíduo não conhece nada sobre o Objeto não há o que possa ser explicitado sobre ele que irá ser apreendido.

Desse modo, embasados pela teoria peirceana, acreditamos que profissionais de informação criam hábitos de agir que guiam a sua leitura profissional, levando a que se utilizem de estratégias de leitura cognitivas e metacognitivas para identificação e seleção de assuntos que tenham pertinência quanto ao texto e sejam potencialmente informativos aos usuários.

4 Metodologia

Para a realização deste estudo foi utilizada a técnica introspectiva de coleta de dados do Protocolo Verbal Individual durante a tarefa de catalogação de assuntos de livros.

A técnica de coleta de dados introspectiva denominada "Protocolo Verbal" consiste na gravação da exteriorização verbal do pensamento de um ou mais indivíduos durante a realização de uma tarefa. O *pensar alto* do sujeito é gravado e transcrito literalmente. No caso do leitor, ele pode exteriorizar seus processos mentais enquanto a informação processada está sob o foco de sua atenção. Ele lê e interpreta ao mesmo tempo, exteriorizando em voz alta tudo o que passa pela sua cabeça durante a leitura (FUJITA, NARDI, FAGUNDES, 2003).

O Protocolo Verbal Individual é uma metodologia útil para a consecução do objetivo desta análise por se tratar de uma técnica introspectiva que permite a observação do processo de leitura, porque o leitor verbaliza o conhecimento processual que possui para o desenvolvimento da atividade, realizando uma leitura consciente. Como também foi exposto por Tartarotti, Dal'Evedove e Fujita (2017, p. 44)

> A introspecção é considerada um evento mental, por definição. Ato pelo qual o sujeito observa os conteúdos de seus processos mentais, assumindo consciência deles. Muitos são os conteúdos mentais passíveis de introspecção, dentre os quais, e de interesse para a pesquisa, figuram o conteúdo do pensamento (conceitos, raciocínios, associações de ideias etc.).

Os dados utilizados para a análise são provenientes de coletas de dados realizadas por pesquisadores participantes do projeto de pesquisa "O contexto da leitura documentária de indexadores de bibliotecas universitárias em perspectiva sócio-cognitiva para a investigação de estratégias de ensino", coordenado pela Profa. Mariângela Spotti Lopes Fujita no período de 2007-2010. Estes dados foram obtidos com o consentimento da coordenadora do projeto e foram analisados com base no interesse desta pesquisa.

As coletas de dados foram realizadas na Rede UNESP de bibliotecas com os catalogadores responsáveis pelas as atividades de catalogação e indexação. Para melhor aproveitamento dos dados, foram selecionadas as coletas referentes à tarefa de Catalogação Original, quando não há registro bibliográfico daquela obra na base de dados bibliográficos ATHENA ou

em bases servidoras, deste modo, o catalogador deverá fazer a catalogação em primeira mão e, portanto, a leitura para análise de assunto da obra.

A coleta de dados com Protocolo Verbal Individual abrange três grupos de procedimentos:

1) **Procedimentos anteriores à coleta de dados**

   a. Definição do universo da pesquisa;

   b. Seleção do Texto-Base;

   c. Definição da tarefa;

   d. Seleção dos sujeitos;

   e. Conversa informal com os sujeitos;

   f. Familiarização com a tarefa do "*Think Aloud*".

2) **Procedimentos durante a coleta de dados**

   a. Gravação do "Pensar Alto" durante a leitura do texto-base.

   b. Entrevista retrospectiva (optativa).

3) **Procedimentos posteriores à coleta de dados**

   a. Transcrição literal das gravações das falas dos sujeitos (Protocolo Verbal Individual);

   b. Leitura detalhada dos dados em busca de fenômenos significativos e recorrentes para construir categorias de análise;

   c. Construção das categorias;

   d. Volta aos dados para retirar trechos da discussão que exemplifiquem cada fenômeno, cada categoria.

Por se tratar de uma pesquisa com reutilização de dados coletados em âmbito de outro estudo, os procedimentos realizados tratam-se dos itens "c" e "d" dos procedimentos posteriores à coleta de dados.

As categorias de análise, selecionadas com base nos estudos teóricos, foram:

- Semiose – Interpretação
- Níveis de interpretante
- Experiência colateral
- Estado de dúvida/crença
- Hábito/mudança de hábito

Em seguida passaremos a exposição dos resultados de nossa análise.

5 Compreensão da leitura documental a partir da semiótica peirceana: resultados e discussões

Em seguida iremos explorar algumas das intersecções possíveis entre a teoria semiótica proposta por Peirce e a leitura documental em Ciência da Informação, com base na análise dos protocolos verbais. Apresentaremos nossos resultados nas categorias de análises definidas previamente.

- Semiose

Para a realização da indexação é necessário que o profissional da informação compreenda o texto que está sob seu exame, pois, só assim a identificação e seleção de conceitos serão bem feitas. Não será possível tornar explícito o conteúdo do texto sem compreensão do mesmo por meio da leitura (FUJITA, 2003).

Catalogador:

"Essa formação nesse sentido contempla compreensão de que a decisão de que a decisão sobre os fundos das escolas é coletiva e é responsabilidade de todos, então "educação administração", "administração escolar", "formação de professores" e "gestores educacionais". Gestores administradores primeiro"

No trecho destacado da fala do sujeito, percebe-se que o bibliotecário catalogador enquanto lê o texto, procede à identificação dos conceitos que considera importantes para o conteúdo do livro, destacadas entre aspas, inclusive estabelecendo uma ordem de prioridade. Este procedimento é feito pela interpretação do conteúdo lido do documento.

Aludindo à teoria semiótica de Peirce é possível compreender o processo de desconstrução do texto inicial, ou de interpretação, como o processo de geração de significado pela mente, por meio do qual o signo é construído. A esse processo Peirce nomeia como "semiose" o processo de construir significados por meio de signos.

Nesse sentido, o próprio texto analisado, o documento que será indexado, é um signo. É uma composição de significado construída por alguém, o autor, para representar uma parcela do mundo que por ele foi interpretada, e que irá, da mesma forma, gerar uma nova composição de significado, na mente do leitor, distinta da primeira, que gerou o documento. Essa nova composição será apenas o efeito produzido pelo signo na mente intérprete, um novo interpretante.

Toda interpretação é, assim, um tipo de modificação do *representamen* observado em um novo interpretante. Podemos considerar que a interpretação é dependente das características emocionais, contextuais, de limitações cognitivas e de objetivos de leitura. Não há garantias de que dois leitores, diante de um mesmo texto, obterão a mesma interpretação, como afirma também Silveira (2007, p. 50) "nada garante que haverá um dia uma interpretação definitiva, nem que somente venha a haver uma única interpretação" (2007, p. 50).

– Níveis de interpretante

É na constatação das diferentes possibilidades interpretativas que reside o principal problema da leitura documental. Como interpretar o texto de um documento para levantar o seu conteúdo principal de forma que este conteúdo, ou o signo que o representará, seja indicador ao usuário do assunto do documento, ou em outras palavras, como agir de modo que a interpretação do profissional de informação assemelhe-se àquela que o próprio usuário faria.

Catalogador

"Eu tenho que ver que público que vai usar, quem vai ser o consumidor desse material"

"então a gente já tem, procura indicar alguma coisa que o nosso usuário tenha familiaridade"

"Um exemplo assim típico disso é "educação matemática" que dentro da pedagogia tem um professor que estuda educação matemática"

"Mas vê-se na área que é uma forma usual de chamar. Então o catálogo tem que se adequar a terminologia e não o pesquisador tem que se adequar ao cabeçalho, porque o objetivo dele é chegar ao material."

Nos trechos destacados, podemos perceber que o bibliotecário catalogador procura se atentar para a demanda possível que o livro catalogado irá atender, tentando compreender o que será de interesse para o usuário, inclusive recobrando exemplos passados de interação com a comunidade.

O leitor profissional busca em sua estrutura cognitiva ferramentas para ajudá-lo na compreensão dos textos que se colocam sob sua análise, de modo que possa gerar outro texto em sua mente, decorrente da interpretação do primeiro texto. Esse produto gerado em sua mente será, no processo de indexação, reavaliado pelo profissional de informação e o produto dessa nova análise, traduzido em termos de uma linguagem de indexação, será o descritor de assunto para o documento.

– Experiência colateral

Catalogador

"Para esse assunto ele manda ver matemática estudo ensino, mas se eu fosse classificar em matemática estudo ensino seria 51-07 pela CDU e não e isso que ele quer, ele quer a educação matemática ele quer Educação"

"Organização escolar é uma remissiva no BIBLIODATA. Então ele vai mostrar qual a forma adotada. Ele manda ver "Escolas organização e administração", tudo bem, então toda vez que eu usar "Organização escolar", que é a forma que eu chamei, ele manda ver "Escolas organização e administração.""

"Eu atualizo o meu, ((ASD)) Gestão educacional. Porque quem está escrevendo é que pode determinar a maneira como os pares/ Gestão educacional ele sugere. [...] Gestão escolar. Qual seria a diferença? Não sei."

> *"Gestão escolar, é a maneira como eles se comunicam, que eu não estou inventando nada, estou seguindo o que outras bibliotecas fazem. Agora em outros casos a gente pega a gente não estaria inventando a gente estaria acatando a opinião de especialistas"*

Podemos observar nos trechos em destaque que o bibliotecário catalogador se apoia em seu conhecimento da linguagem, conhecimento prévio profissional, para determinar se a interpretação que fez do documento é relevante e pode relacionar-se com a demanda usuária. O catalogador também manifesta não ter certeza quanto aos termos que melhor representam os conceitos identificados no documento, buscando a terminologia aplicada em outros catálogos.

Quando dissemos que o indexador desconhece a área de assunto especializada, em termos semióticos, o que estamos a sugerir é que o profissional desconhece o referente ao qual o signo procura representar.

Embora tenha contato com o signo, não teve a experiência colateral com o Objeto Dinâmico, desconhece suas propriedades e suas idiossincrasias. Se aceitamos que a compreensão necessita de conhecimentos prévios do intérprete para se realizar, entendemos que a semiose proveniente dessa relação apresentará limites, pois o indexador não tem conhecimento prévio da área específica tal como possui o usuário especialista no domínio.

Souza e Almeida (2012) lembram que a experiência colateral do indexador pode ir sendo construída com o tempo, na forma de experiência profissional, utilizando-se de linguagens de indexação como referentes, processo observado no trecho destacado do Protocolo Verbal. A linguagem de indexação, instrumento que, se propõe a conhecer e tornar explícitos os significados atribuídos a determinados signos em um domínio, servirá ao indexador de contexto para a interpretação do signo/documento e, dessa forma, apenas constrói uma sugestão de referente.

– Estado de dúvida/crença

Compreendemos que o leitor profissional passa do estado de dúvida ao estado de crença constantemente durante a leitura, e para deixar o estado de dúvida o profissional aciona uma ou mais das estratégias de

leitura mencionadas anteriormente, no intuito de solucionar a dificuldade encontrada e retornar ao estado de crença.

> **Catalogador**
>
> *"Agora vou ler a última capa para ver porque aqui tem mais a ver com formação de professores e administração escolar do que com história."*
>
> *"olha "administração escolar" a palavra bastante repetida"*
>
> *"esse livro, vamos ver ele já está falando do objetivo deles (~~~) discorre sobre gestão pedagógica"*

No trecho destacado é possível observar que o bibliotecário catalogador faz uso de estratégias de leitura para atingir o objetivo da indexação, ao direcionar sua atenção a um determinado elemento do livro, no caso a última capa, ao buscar a frequência em que determinadas palavras são usadas no texto, assim como, quando aumenta seu nível de atenção ao perceber que se trata de uma passagem relevante para a compreensão do assunto do livro, no caso o objetivo do estudo.

A alternância entre o estado de dúvida e a crença, alcançada por meio de uma determinada ação de pensamento, leva o bibliotecário a ter esta ação como um procedimento que pode ser repetido quando em situação de dúvida semelhante, o que Peirce denominou como a construção de um hábito de agir.

As estratégias que mencionamos têm por objetivo, em uma análise semiótica, colaborar com o indexador para que sua interpretação construa um interpretante que se assemelhe àquele que o usuário poderia alcançar, dado que em razão de seu conhecimento prévio da área temática, o usuário possui condições interpretativas diversas das que possui o indexador.

As estratégias são potencialmente conscientes e, em geral se direcionam a solução de um problema. Tais estratégias podem oferecer ao indexador a perspectiva do sistema de informação, uma forma de direcionar sua leitura para a busca de uma interpretação coerente com a demanda, condições para monitorar se a leitura está atingindo seu objetivo e se há compreensão, conhecimento e segmentação do texto de modo a atentar para áreas e aspectos principais e, para a correção de suas ações, se necessário (NEVES; 2011; REDÍGOLO; FUJITA, 2015).

Observando as estratégias de leitura mencionas sob a ótica da filosofia peirceana, podemos compreendê-las enquanto hábitos de agir adquiridos para conduzir a leitura profissional. Hábitos em Peirce são tendências inatas ou adquiridas para ter condutas semelhantes diante de situações semelhantes, em outras palavras "O que nos leva a, dadas certas premissas, retirar esta ou aquela inferência é uma tendência de espírito, seja ela constitucional ou adquirida" (PEIRCE, 1975, p. 74).

– Hábito/mudança de hábito

As reflexões feitas no âmbito da categoria anteriormente exposta estão fortemente relacionadas com esta categoria. Aqui, pretendemos demonstrar como um hábito pode ser determinado, não a partir da experiência empírica, mas incutidos no indivíduo por influência externa, por exemplo, a partir de procedimentos padrão que devem ser adotados.

Catalogador

"Agora eu vou olhar todos os campos para ver se está de acordo com os padrões da UNESP, então algumas coisas a gente fez tanto que já decorou, né?"

"Isso eu faço com um roteirinho do lado, pois a gente esquece muito"

No trecho destacado é possível observar que o bibliotecário catalogador procura se apoiar em seus hábitos de agir, inclusive demonstrando insegurança quando não pode recorrer a um hábito previamente construído. Porém, também podemos notar o processo de construção de um hábito de forma indireta, não pelo exercício do pensamento individual, mas por meio de um procedimento padronizado para uma dada atividade.

6 Considerações finais

Na cadeia de processos documentais realizados em unidades de informação, a representação do conteúdo de documentos reveste-se de importância essencial, dada a sua função de mediar a relação entre os usuários e a informação documental. A representação da informação, em outras palavras, objetiva que a informação possa ser corretamente recuperada, acessada e, por fim, utilizada. Esse entendimento confere importância cabal

aos estudos de leitura para indexação no âmbito da organização do conhecimento, buscando compreender as ações realizadas por profissionais de informação, em seu contexto social e cognitivo, a formação que possuem para realizar estas atividades e a sistematização da mesma, materializada em políticas de tratamento da informação e manuais de atuação.

As teorias filosóficas de Charles Sanders Peirce surgem nesse panorama como uma possibilidade para ampliar os horizontes teóricos da Ciência da Informação, em especial, da área denominada de organização da informação e do conhecimento.

O processo de indexação, em especial, constitui-se de interpretação e de representação, conceitos essenciais na teoria semiótica proposta por Peirce, o que leva a assegurar que o conhecimento de tal teoria ofereça aos indexadores uma perspectiva ampliada de sua atividade, de modo a compreendê-la melhor.

Para realizar a indexação é preciso que haja compreensão do texto abordado, sem a qual não se torna possível a identificação de conceitos de relevância presentes no texto, dessa forma, mesmo a leitura documental é um processo interpretativo e, portanto, onde ocorre semiose. O nível de compreensão, entretanto, dificilmente será aprofundado, permanecendo numa apreensão superficial, construindo interpretantes superficiais em relação ao objeto, dado a inexistência de experiência colateral com o Objeto Dinâmico do signo.

Ainda assim, é possível observar que a semiose não ruma sempre a um nível de interpretante mais desenvolvido, podendo caminhar a níveis menos desenvolvidos, no caso da indexação é necessário que assim seja para construção da informação documental. Há, portanto, uma perda quantitativa de significado, bem como uma perda qualitativa, passando de conteúdo simbólico, para um nível indicial, pois, a linguagem documental delimita o significado dos termos, de modo a torná-los unívocos.

O estudo da compreensão peirceana de hábito e mudança de hábitos também pode engrandecer e dar consistência aos estudos de leitura documental ao oferecer novas possibilidades de compreensão do contexto cognitivo do leitor profissional, permitindo que se ampliem as possibilida-

des de ensino desses procedimentos a indexadores iniciantes, bem como, em treinamentos para atualização profissional e/ou formação em serviço.

O uso da metodologia introspectiva do Protocolo Verbal Individual se mostrou relevante para observar os interesses deste estudo tendo explicitado fenômenos semióticos que possivelmente não poderiam ser observados por outras metodologias não introspectivas.

Como visto, as dificuldades encontradas para a análise de assunto são muitas e as possibilidades de alcançar um nível ideal de representação de assunto são limitadas pelas múltiplas possibilidades interpretativas presentes na relação leitor-texto. Entendemos, então, que a indexação deve estar fortemente embasada nas concepções orientada ao conteúdo e à demanda e a unidade de informação deve conhecer o seu público potencial, por meio de estudos de comunidades, estudos de usuários, análise de domínios e outros meios, de modo que possa estar o mais claro possível ao indexador quais são os interesses informacionais deste público.

Referências

ALMEIDA, C. C. ; FUJITA, M. S. L.; REIS, D. M. Peircean semiotics and subject indexing: contributions of speculative Grammar and pure logic. *Knowledge Organization*, v. 40, n. 4, p. 225-241, 2013.

DAL' EVEDOVE, P. R.; ALMEIDA, C. C.; FUJITA, M. S. L. Interpretação de conceitos no processo de indexação: aspectos e reflexões semióticas. In: RIBEIRO, F; CERVEIRA, M. E. (Orgs.). *Informação e/ou conhecimento*: as duas faces de Jano. Porto: Faculdade de Letras da Universidade do Porto - CETAC. MEDIA, 2013. p. 190-203.

DUMONT, L. M. M. Os sentidos da leitura e a subjetividade. In: NAVES, M. M. L.; KURAMOTO, H. *Organização da informação*: princípios e tendências. Brasília-DF: Briquet de Lemos/Livros, 2006.

FUJITA, M. S. L. A identificação de conceitos no processo de análise de assunto para indexação. *Revista Digital de Biblioteconomia e Ciência da Informação*. v. 1, n. 1, 2003.

FUJITA, M. S. L. A leitura documentária na perspectiva de suas variáveis: leitor-texto-contexto. *Datagramazero*, Rio de Janeiro, v. 5, n. 4, 2004. Disponível em: <http://dgz. org. br/ago04/F_I_art. htm>.

FUJITA, M. S. L. A representação documentária no processo de indexação com o modelo de leitura documentária para textos científicos e livros: uma abordagem cognitiva com protocolo verbal. *Ponto de Acesso*, Salvador, v. 7, n. 1, p. 42-66, abr. 2013. Disponível em: <www.pontodeacesso.ici.ufba.br>.

FUJITA, M. S. L.; NARDI, M. I. A.; FAGUNDES, S. A. A observação da leitura documentária por meio de Protocolo Verbal. In: RODRIGUES, G. M.; LOPES, I. L. (Orgs.). *Organização e representação do conhecimento na perspectiva da Ciência da Informação*. Brasília: Thesaurus, 2003. p. 141-178.

HJØRLAND, B. Toward a theory of aboutness, subject, topicality, theme, domain, field, content… and relevance. *Journal of the American Society for Information Science and Technology*, v. 52, n. , 2001, 774-778.

LANCASTER, F. W. *Indexação e resumos*. Brasília: Briquet de Lemos, 2004.

LARA, M. L. G. Algumas contribuições da semiologia e da semiótica para a análise das linguagens documentárias. *Ciência da Informação*, Brasília, v. 22, n. 3, p. 223-226, set/dez. 1993.

LARA, M. L. G. Linguagem documentária e terminologia. *Transinformação*, Campinas, v. 16, n. 3, p. 231-240, 2004.

MAI, J-E. Semiotics and indexing: an analysis of the subject indexing process. *Journal of Documentation*. v. 57, n. 5, (September 2001), p. 591-522.

MOURA, M. A. Leitor-bibliotecário: interpretação, memória e as contradições da subjetividade em processos de representação informacional. In: NAVES, M. M. L.; KURAMOTO, H. *Organização da informação*: princípios e tendências. Brasília-DF: Briquet de Lemos/Livros, 2006.

NEVES, D. A. B. Leitura e metacognição: uma experiência em sala de aula. *Encontros Bibli*: Revista Eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, v. 12, n. 24, p. 1-9, 2007.

NEVES, D. A. B.; RAMALHO, F. A.; FIGUEIREDO, H. A. *O uso de estratégias metacognitivas pelos alunos de graduação do Curso de Biblioteconomia da Universidade Federal da Paraíba/Brasil*. In: Formación, investigación y mercado laboral en Información y Documentación en España y Portugal: Formação, investigação e mercado de trabalho em informaçao e documantaçao en Espanha e Portugal: [III Encuentro Ibérico de Docentes e Investigadores en Información y Documentación]. 2008. p. 329-338.

PEIRCE, C. S. *Semiótica e filosofia*. São Paulo: Cultrix, 1975.

PEIRCE, C. S. Uma visão do Pragmaticismo. In: FREGE, G.; PEIRCE, C. S. *Os Pensadores*. São Paulo: Abril Cultural, 1980.

PEIRCE, C. S. *Semiótica*. São Paulo: Perspectiva, 2010.

REDIGOLO, F. M. *O uso de linguagens documentárias por catalogadores de assunto em contexto de Bibliotecas Universitárias*: uma abordagem sociocognitiva com protocolo verbal. 2007. 87 f. Relatório parcial período de agosto a dezembro 2007; CNPq/PIBIC. Marília, 2007.

REDIGOLO, F. M.; ALMEIDA, C. C. Algumas contribuições da perspectiva filosófico-semiótica de Peirce para a análise de assunto. *DataGramaZero*, Rio de Janeiro, v. 13, n. 3, jun. 2012.

REDÍGOLO, F. M.; FUJITA, M. S. L. A leitura profissional do catalogador e seu papel como mediadora da informação. *Informação e Informação*, Londrina, v. 20, n. 3, p. 356 - 376, set./dez. 2015.

SILVEIRA, L. F. B. *Curso de semiótica geral*. São Paulo: Quartier Latin, 2007.

SOUZA, B. P.; ALMEIDA, C.C. de. Um olhar semiótico sobre o processo de indexação: a questão da representação e do referente. *Informação & Sociedade*: Estudos, João Pessoa, v. 22, n.2, p. 23-34, (maio/ago. 2012).

TARTAROTTI; R. C. D.; DAL'EVEDOVE, P. R.; FUJITA, M. S. L. Protocolo Verbal em Grupo e a pesquisa brasileira em Organização e Representação do Conhecimento. *Encontros Bibli*: revista eletrônica de biblioteconomia e ciência da informação, v. 22, n. 48, p. 41-58, jan./abr., 2017.

Contribuições teóricas da Semiótica para a leitura documentária

Daniela Majorie Akama dos Reis

Introdução

A leitura é um processo ativo, no qual o leitor traz toda uma vida de experiências para o texto e utiliza essa experiência para interpretar e elaborar sobre seus conteúdos. Os escritores confiam que os leitores farão inferências que lhes permitirão evitar ter de contar tudo em detalhes entediantes (ELLIS, 1995, p. 62).

Em Organização do Conhecimento[1], a leitura é a forma pela qual os profissionais iniciam diversos processos, como por exemplo, a análise de assunto e indexação. A leitura com objetivos documentários, realizada por um profissional nunca ocorrerá de maneira determinada. Cada mente funciona de forma única, fatores como conhecimento prévio, domínio em que o sujeito atua (e/ou atuou anteriormente), educação acadêmica, convívio com outras pessoas, tudo, influencia na forma como tal pessoa interpreta determinada informação. Consequentemente, é fundamental revisar e atualizar os rumos dos constantes estudos sobre as variáveis de leitura, o que justifica este estudo.

[1] [...] "Organização da informação em registros bibliográficos, incluindo índices de citação, texto completo e internet" (HJØRLAND, 2003, p. 1)

Com base em pesquisas anteriores, considera-se relevante citar[2] algumas das diversas variáveis que envolvem a leitura, e consequentemente, a leitura documentária: o **domínio**, contemplado do ponto de vista tanto do documento quanto do leitor; **conhecimento prévio** dos profissionais; **tipo de documento** (pensando em estruturas textuais de documentos e suas respectivas áreas – considerando a questão do domínio do acervo em geral); entre outras variáveis relacionadas ao profissional como leitor (ambiente de trabalho, pressão, aspectos não mensuráveis).

Hjørland considera que os métodos de Organização do Conhecimento em Biblioteconomia e Ciência da Informação são, no nível mais profundo, baseados nas mesmas suposições filosóficas dos métodos de ciência e erudição. Isto implica em que a discussão fundamental da base da Organização do Conhecimento é fortemente conectada à discussão de diferentes teorias em epistemologia (2003, p. 9).

A **Semiótica**, ou Teoria Geral dos Signos, é uma indagação sobre a natureza dos signos e suas relações, entendendo-se por **signo** tudo aquilo que represente ou substitua alguma coisa, em certa medida e para certos efeitos (PIGNATARI, 2004, p.21). De acordo com o autor, entende-se que "toda e qualquer coisa que se organize ou tenda a organizar-se sob a forma de linguagem, verbal ou não, é objeto de estudo da semiótica" (PIGNATARI, 2004, p.15).

A leitura documentária, portanto, independentemente de seus objetivos e produtos, é passível de análise por meio de aspectos manifestados na teoria semiótica, pois um texto "não apresenta significado nele mesmo, é necessário que haja uma interação a partir da leitura para que o significado e conceitos que o texto pretende passar sejam compreendidos pelo leitor" (MAI, 2004, p. 604).

Em dissertação de mestrado, defendida por Reis em 2012, foram coletados Protocolos Verbais Individuais - PVIs[3] nas três áreas do conheci-

---

[2] Não é o foco desta pesquisa apresentar detalhadamente cada uma delas.

[3] Consiste na gravação da exteriorização verbal de pensamento durante a atividade de leitura. Isso é possível porque o leitor pode exteriorizar seus processos mentais enquanto a informação processada está sob o foco de sua atenção. Referem-se ao "Think Aloud" ("Pensar Alto"), em que o indivíduo lê e interpreta ao mesmo tempo, exteriorizando em voz alta tudo o que "passa pela sua cabeça" durante a leitura. [...] o "Pensar alto" do informante é gravado e transcrito literalmente, produzindo protocolos verbais. Protocolos são geralmente definidos como relatos verbais dos processos mentais conscientes do informante (FUJITA, NARDI e FAGUNDES, 2003, p. 142).

mento (Humanas, Exatas e Biológicas), em bibliotecas universitárias, durante a catalogação de assunto de livros científicos. Concluiu-se por meio de análise detalhada da observação de estruturas textuais durante leitura, que seria necessária outra forma de analise com base em critérios mais evoluídos (já considerando resultados anteriores) para entender como os profissionais agem durante o processo interpretativo com fins documentários.

Em artigo publicado em 2013, Almeida, Fujita e Reis consideram "a abordagem semiótica do processo de indexação de assunto como um campo promissor para os estudos da Organização e Representação do Conhecimento" (2013, p. 239). Pesquisas que relacionam teorias da semiótica tem sua importância no tratamento temático da informação, principalmente devido à carência de estudos à esta abordagem específica. De acordo com os autores, estudos com base nos três tipos de raciocínio (Abdução, Dedução e Indução), explicados por Peirce, podem esclarecer de forma mais genuína o processo de leitura documentária:

> Em síntese, com a abdução sustentamos os processos de leitura (percepção) e criação de sugestões hipotéticas de representação do assunto; cumpre a dedução tratar da seleção de termos de representação segundo a generalização intelectual das consequências e; por último, resta a indução ancorar as funções do teste e comparação dos potenciais assuntos com às linguagens de indexação, linguagem do usuário e linguagem do sistema de informação, com o objetivo de continuamente avaliar e aprimorar as representações de assunto (ALMEIDA; FUJITA; REIS, 2013, p. 240).

Portanto, propõe-se um estudo sobre a leitura documentária e aspectos teóricos conhecidos na semiótica - especificamente, a Abdução, Dedução e Indução - e seus efeitos na interpretação de profissionais responsáveis pela leitura documentária de livros, fundamentada em literatura em Organização do Conhecimento e semiótica – principalmente Peirce e autores baseados na teoria de Peirce.

Tem-se como objetivo deste estudo teórico, alicerçar pesquisas aplicadas dentro da temática da leitura documentária sob um ponto de vista de teorias semióticas.

Serão apresentados conceitos, com base em pesquisas anteriores, e em literatura nacional e internacional, sobre a leitura, leitura documen-

tária e por fim, uma sincronização com aspectos da teoria semiótica. A primeira seção tratará da semiótica e perspectivas para pensar na leitura, seguida de uma seção sobre leitura documentária, uma discussão dos temas tratados e resultados, e por fim, as considerações finais e sugestão para pesquisas aplicadas.

Aspectos semióticos da leitura

Para Eco, a semiótica está relacionada a tudo que possa ser **assumido** como signo. "É signo tudo quanto possa ser assumido como um substituto significante de outra coisa qualquer. Esta outra coisa qualquer não precisa necessariamente existir, nem subsistir de fato no momento em que o signo ocupa seu lugar" (2014, p. 4).

Almeida (2012, p. 50), explica a Semiótica peirceana, como uma ciência formal, que está subdividida nos ramos: Gramática Especulativa, Lógica Pura e Retórica Especulativa.

A lógica é o ramo da filosofia que permite uma análise do pensamento. Consiste numa operação mental que possibilita, através do raciocínio, o surgimento de novas proposições através de proposições já existentes (DIAS; NAVES, 2013, p. 83).

Considera-se inerente a relação da lógica com a leitura, independente do objetivo desta última (documentária ou fluida), pois esta demanda habilidades e estratégias específicas por parte dos leitores.

Diante da análise da concepção de lógica de Peirce, torna-se evidente que o esforço necessário para a análise de assunto está relacionado à *lógica docens*[4], pela exigência da capacidade inventiva, generalização e teoria, por parte do indexador, muito além da lógica do bom senso do homem comum (DIAS; NAVES, 2013, p. 85).

Almeida, Fujita e Reis (2013, p. 240) consideram "fundamental destacar o papel da Lógica como matriz teórica para conceber o processo de indexação como inferencial". Pode-se concluir, que à lógica, é atribuída essencial importância quando pretende-se investigar aspectos inferenciais da leitura documentária. Ainda, "a Lógica Pura, segundo

[4] Resultado de estudo científico, conhecido como "logica docens".

ramo da semiótica de Peirce, procura responder a relação dos signos com os objetos"

Neris (2006, p. 3) faz considerações a respeito da leitura do ponto de vista da teoria semiótica:

> [...] é essencialmente uma semiose, uma atividade primordial cujo resultado é correlacionar um conteúdo a uma expressão dada e transformar uma cadeia de expressão em uma sintagmática de signos. Tal performance pressupõe uma **competência** do leitor, comparável, ainda que não necessariamente idêntica, à do produtor do texto. Se, no momento da leitura normal, o fazer receptivo e interpretativo do enunciatário-leitor continua implícito, sua explicitação, sob forma de procedimentos de análise estabelecidos, tendo em vista a reconstrução do sentido, constitui tarefa da semiótica (grifo nosso).

Para Eco (2014, p. 60), "A semiose explica-se por si só. Esta contínua circularidade é a condição normal da significação, e é isto que permite o uso comunicativo dos signos para referir-se a coisas".

Almeida, Fujita e Reis (2013), exploram o modelo semiótico da indexação de Mai (2001)[5], e explicam a semiose no processo de indexação, que seria entendida como "uma sequência de atos interpretativos geradores de signos a partir de seus passos e elementos" (2013, p. 237). Na visão dos autores, existe a possibilidade da continuidade da semiose no momento da busca realizada pelos usuários, ou seja, quando estes entram em contato com os termos selecionados pelo bibliotecário.

Pode-se constatar que o sujeito que realiza a leitura documentaria em bibliotecas – preferencialmente – será um bibliotecário munido de conhecimento prévio e de diversas estratégias profissionais de leitura. Este conjunto de fatores torna o bibliotecário competente para a leitura com fins de sumarização, e de maneira inferencial, diferenciando-o de um leitor comum, que realiza uma leitura espontânea e do documento na íntegra – fora do domínio profissional.

Ao contemplar a leitura sob um viés da semiótica, pode-se avaliar que "ler é produzir um texto que tem como referente outro texto. A leitura

---

[5] MAI, J-E. Semiotics and indexing: na analysis of the subject indexing process. *Journal of Documentation,* v. 57, n. 5, p. 591-622, 2001.

e a interpretação são constituídas por uma relação intertextual, ou seja, um texto que remete a outro, que é a sua razão de ser" (NERIS, 2006, p. 5). Observa-se aqui, a influência do conhecimento prévio do leitor em determinada área, quanto mais este conhece sobre tal assunto, melhor e maior será sua habilidade de criação de hipóteses por meio da abdução.

Thellefsen (2002, p. 76) também explica que em alguns casos, o contexto é um importante elemento a ser considerado. "Buscamos um contexto, que é tão bem definido que pode proporcionar inúmeros signos com significado específico". Entende-se que o que Thellefsen considera em artigo de 2002, é o contexto como fator decisivo na interpretação de signos, considera-se aqui o mesmo para profissionais – bibliotecários – que realizam a leitura documentária. O conhecimento prévio sobre um assunto dentro um domínio especifico é o fator que destaca um profissional de outros.

Na seção seguinte, serão apresentados conceitos, características e tipos de estratégias da leitura e da leitura documentária.

A leitura documentária

Pinto e Gálvez (1999, p. 40) definem a leitura de maneira geral como passo inicial para a aquisição de informações textuais. É realizada de forma automática, quase que inconscientemente, por estar muito enraizada em nossa vida diária.

A leitura, na concepção de Fujita (2004, p. 2):

> [...] apesar da individualidade do ato realizado, é um ato social porque existe um processo de comunicação e de interação entre o leitor e o autor do texto, ambos com objetivos estabelecidos anteriormente dentro do contexto de cada um. Apesar de, aparentemente simples e tão natural, o processo de leitura possui uma complexidade que está subjacente porque depende do processamento humano de informações e da cognição de quem lê, de um texto elaborado por um autor e do contexto de ambos, o que determina os objetivos da leitura.

Segundo Neves (2007, p.2) "no que diz respeito à compreensão de um enunciado, a leitura de um texto é uma atividade cognitiva que

requer esforço mental/cognitivo, envolvendo: percepção, memória, inferência e dedução".

Para Livingston, o conhecimento sobre as variáveis de estratégias de leitura inclui conhecimento sobre estratégias cognitivas e metacognitivas, bem como conhecimento condicional sobre quando e onde é apropriado usar tais estratégias. (1997, p. 1)

A metacognição é um assunto bastante estudado em pesquisas sobre leitura documentária, e de acordo com Livingston (1997, p. 2) significa: "'pensar sobre o pensamento' e envolve supervisionar se uma meta cognitiva foi cumprida. Este deve ser o critério de definição para determinar o que é metacognitivo".

Deve-se ter em mente que "as estratégias cognitivas e metacognitivas estão estreitamente entrelaçadas e dependentes uma da outra, qualquer tentativa de examinar uma sem reconhecer a outra não forneceria uma situação adequada" (LIVINGSTON, 1997, p. 2).

Sobre a leitura, conclui-se que é inerente ao ser humano o uso de estratégias cognitivas e metacognitivas, e é por meio destas que o leitor proficiente está apto a formular hipóteses para compreender o texto.

A leitura documentária, por sua vez, apresenta aspectos que diferem da leitura habitual, pois, não é realizada para lazer ou aprendizagem, nem é prazerosa, muito pelo contrário. O alto grau de incerteza, ansiedade e responsabilidade contido na atividade já mostra que ela traz pouca satisfação. (DIAS; NAVES, 2013, p. 41)

Para Dias, Neves e Pinheiro (2006, p. 142), os leitores:

> [...] identificam informações relevantes, lêem as partes aparentemente mais importantes, fazem inferências, lêem em voz alta, repetem e reformulam uma idéia buscando sua correspondência na memória de trabalho. Tomam notas, fazem pausas para refletir sobre o texto, elaboram paráfrases, buscam padrões textuais, fazem predições. Relacionam partes do texto buscando esclarecer dúvidas, interpretam o texto, emitem juízos de valor sobre a qualidade do texto e a veracidade do relato, entre outras.

Tanto a redação quanto a leitura de um texto, são processos repletos de complexidade. Koch (1998, p. 25) afirma que "o sentido não está no texto, mas se constrói a partir dele, no curso de uma interação".

Um texto não apresenta significado nele mesmo, é necessário que haja uma interação a partir da leitura para que o significado e conceitos que o texto pretende passar sejam compreendidos pelo leitor (MAI, 2005, p. 604).

Em PVIs é comum identificar tal comportamento por parte dos leitores proficientes. Estes verbalizam suas estratégias, possibilitando melhor entendimento de como processam a informação durante a leitura documentária.

Assim como na leitura habitual, durante a leitura de um texto com objetivos documentários, "são ativados esquemas variados, desde conhecimento de vocabulário, conhecimento da estrutura textual, do assunto, até conhecimento de mundo" (FUJITA; NARDI; SANTOS, 1998, p. 14). O quadro abaixo apresenta uma comparação de estratégias de leitura em quatro concepções teóricas distintas:

Brown	**Kato**	**Cavalcanti**	**Cintra (citando teóricos da ciência da cognição)**
Skill Estratégia	Estratégia cognitiva Estratégia metacognitiva	Estratégia automática Estratégia controlada	Estratégia automática Estratégia controlada

Quadro 1: Estratégias de leitura conforme as concepções teóricas (p. 19)
Fonte: Fujita, Nardi e Santos (1998)

Nota-se que todas as concepções fazem uma distinção entre o que seriam estratégias cognitivas e o que seriam estratégias metacognitivas. O que Kato considera como estratégia cognitiva (*Skill* para Brown), Cavalcanti e Cintra consideram como estratégia automática; E o que Kato considera como estratégia metacognitiva (Estratégia para Brown), Cavalcanti e Cintra consideram como estratégia controlada. Desta forma, as estratégias cognitivas, ocorrem na leitura fluida, do documento na íntegra, e as estratégias metacognitivas, ocorrem durante uma leitura com objetivos documentários.

Na leitura documentária, "não é necessário, nem aconselhável uma leitura linear, letra por letra, palavra por palavra, o leitor avança no texto à medida que consegue predizer o que vem a seguir" (FUJITA; NARDI; SANTOS, 1998, p. 20). As autoras explicam como o leitor com objetivos documentários deve proceder para efetuar a leitura documentária de forma eficaz. Ele deve:

> [...] buscar detectar a estrutura do texto. O leitor que tem facilidade de reconhecer as superestruturas textuais capta melhor as ideias principais do texto, do que um leitor que lê linearmente, fazendo esforços desnecessários para compreender trechos isolados, mas sem apoiar-se na estrutura textual e seus conhecimentos prévios para inferir significados, e levantar hipóteses que o ajudarão a apreender a temática global (FUJITA; NARDI; SANTOS, 1998, p. 20) (grifo nosso).

A Leitura documentária está presente em diversos processos na Organização do Conhecimento. Foi mencionado anteriormente que existem diversos aspectos, que de maneira previsível, influenciam na leitura documentária – independentemente do processo e dos produtos obtidos por meio do tratamento de informações.

Alguns fatores podem afetar o processo de leitura, e consequentemente, a leitura documentária. Kato (1995) define as condições de leitura: a) o grau de maturidade do sujeito como leitor; b) o nível de complexidade do texto; c) o estilo individual; d) o gênero do texto.

Cintra (1983, p. 5) explica que a análise de documentos com fins documentários (indexação ou catalogação de assunto) pode ser realizada pela leitura do documento processada pelo cérebro humano, ou por máquina. Considera-se que a leitura por uma pessoa seja mais adequada dentro de domínios específicos, devido à bagagem de informação e conhecimento especifico em dada área, adquiridos ao longo da vida (aspectos mais detalhados a seguir).

Para Dias e Naves (2013, p. 39):

> O processo de leitura depende da competência comunicativa do leitor, competência essa que sofre influência de vários fatores e, dentre esses fatores, destacam-se: (a) a ação da memória que, incessantemente, relaciona o não conhecido ao conhecido e (b) a participação da razão e suas atividades complementares de indução e dedução, análise e síntese.

Thellefsen (2002, p. 77) supõe que uma pessoa inserida em um domínio de conhecimento bem definido possui um nível maior de conhecimento em relação aos conceitos do que uma pessoa fora do mesmo. A escolha das áreas das bibliotecas universitárias pesquisadas por Reis (2012) teve como base este argumento. Acredita-se que bibliotecários inseridos em domínios específicos observariam estruturas textuais de forma diferenciada (tendo em mente também, os tipos de estrutura textual dos livros e suas variações dentro de cada área).

Dias e Naves (2013, p. 27) explicam a necessidade da observação de estruturas textuais para a análise de documentos por meio da leitura:

> Para a análise de assunto que se realiza em textos escritos é preciso que seja feita uma leitura que possibilite a extração de conceitos que sintetizem o conteúdo desses textos. Sabe-se da importância do texto, de estruturas e de tipos de textos, bem como a forma específica de leitura que deve ser feita pelo indexador.

Como resultados de pesquisa de 2012, chegou-se a diversas conclusões, por meio da análise dos PVIs durante a leitura documentária. Foi possível observar quais as partes da estrutura textual são mais consultadas para a análise de assunto dos livros por parte dos bibliotecários em cada área (bibliotecas de Humanas, Biológicas e Exatas).

Os profissionais responsáveis pela análise de assunto em bibliotecas na área de Humanas apresentam esquemas de leitura documentária que proporcionam melhores resultados com relação à complexa estrutura dos documentos da área de Humanas, mesmo que estas envolvam a leitura de quase ou todas as partes da estrutura textual do livro.

Por outro lado, nas áreas de Biológicas e Exatas, os livros usados durante as coletas de PVIs apresentam uma estrutura mais uniforme de maneira geral, com um título bem representativo e um sumário que lista grande parte dos conteúdos apresentados nos capítulos. Esses fatores auxiliam o bibliotecário no momento da leitura documentária, mesmo quando os livros estão em outra língua (que o leitor não domina).

Por meio da análise da leitura documentária dos bibliotecários em cada área, foi possível observar as partes da estrutura textual que são

mais usadas para a análise de assunto dos livros. Os resultados foram expostos em quadros, listando as partes das estruturas textuais por área e comparando-as.

Ao considerar as variáveis que envolvem a leitura, e com base em Kato (1995), foi elaborado o seguinte quadro:

As condições de leitura (KATO, 1995)	Leitura documentária (categorização para estudos)
a) o grau de maturidade do sujeito como leitor;	• Conhecimento prévio/ enciclopédico/ de mundo dentro de um domínio específico;
b) o nível de complexidade do texto; d) o gênero do texto;	• Livros em áreas especificas (Humanas) dentro de um domínio específico;
c) o estilo individual;	• Aspectos semióticos da leitura (observação da Abdução, Dedução e Indução como características inferenciais);

Quadro 2: As condições de leitura e as variáveis observadas na Leitura documentária
Fonte: Elaborado pela autora

Embora os resultados da dissertação de 2012 tenham sido significativos – no sentido de que se pode notar a tendência de analise de partes de estruturas textuais de livros em cada área do conhecimento – considera-se necessário pesquisar a leitura de maneira mais aprofundada, origem do interesse pelos aspectos semióticos observados na leitura em pesquisa atual.

Discussão e resultados

Para Eco (2014, p.118), logicamente falando, "**interpretação** é uma INFERÊNCIA. Inicialmente, assemelha-se ao tipo de inferência lógica que Peirce chamou de 'abdução' (e em certos casos de 'hipótese')" (grifo nosso).

Diversos autores explicam a abdução, dedução e indução. Eco (2001) apresenta os conceitos com base na teoria de Peirce. De acordo com o proposto nesta pesquisa, daremos mais enfoque à abdução:

> A abdução ou hipótese é amplamente descrita por Peirce em diversos pontos da sua obra [...][6]. Comparada com a dedução e a indução, ela dá lugar aos três diferentes esquemas inferenciais da figura seguinte onde as

[6] [cf. em particular 1902b, tradução italiana, p. 105-6; 1878, ed. 1951-58, §§ 619-635]

casas delineadas a cheio exprimem os estádios argumentativos para os quais se têm proposições já verificadas e as casas a tracejado os estádios argumentativos produzidos pelo raciocínio (ECO, 2001, p. 54):

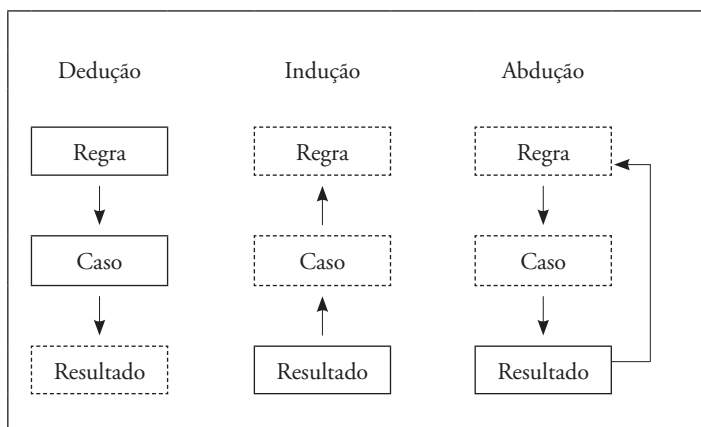



Figura 1: Dedução, Indução e Abdução de acordo com ECO
Fonte: Eco, 2001, p. 54

A Dedução prova que algo deve ser; A indução mostra que algo está operacional; Abdução simplesmente sugere que algo pode ser (PEIRCE, 1931-1958, 171).

Para Eco:

> Que depois haja indícios circunstanciais e contextuais a dirigir-me para o estabelecimento da regra, não altera em princípio a estrutura do processo interpretativo. A abdução intervém também quando devo interpretar figuras retóricas e quando devo interpretar vestígios, sintomas, indícios (veja-se a referência hipocrática ao contexto). O mesmo acontece, no entanto, quando quero interpretar o valor que um dado enunciado, uma palavra-chave, todo um episódio, assumem num texto. A Abdução representa, pois, o propósito, a tentativa arriscada, de um sistema de regras de significação à luz das quais um signo adquirirá o seu significado (2001, p. 56)

O fato da abdução "simplesmente **sugerir** que algo pode ser", coloca a responsabilidade da interpretação nas estratégias metacognitivas desenvolvidas pelo profissional durante toda sua vivencia profissional dentro de um ou vários domínios.

Seguindo essa linha de pensamento, na leitura documentária para indexação, Almeida, Fujita e Reis explicam que "a abdução é a etapa criativa na indexação, a dedução é a etapa generalizadora e a indução é a fase do teste e da continuidade da representação dos assuntos" (2013, p. 238).

Em diversos processos que envolvem a leitura documentária (indexação, catalogação de assunto e classificação), esta, é sempre o primeiro contato do profissional com o documento. É neste momento que se pode observar a atividade inferencial e/ou abdução.

Desta forma, considera-se a abdução, como etapa mais importante observada no processo de tratamento temático da informação, pois é neste momento que o bibliotecário exibe toda sua peculiaridade interpretativa por meio da leitura documentária.

Em etapas seguintes de tais processos de tratamento temático, destaca-se a importância do uso de manuais, ferramentas, políticas, etc. Durante a abdução, o leitor proficiente tem a liberdade para refletir sobre o assunto do documento, antes de traduzir tudo que ele considera importante na representação para uma linguagem documentária, observa-se neste momento, a importância desta etapa inicial. As etapas finais não poderão ocorrer de forma apropriada, sem a devida atenção na etapa inicial – leitura documentária.

À vista disso, podemos considerar que a abdução ocorre durante o primeiro contato do sujeito com o texto, momento em que este aplica todo seu conhecimento prévio, e simultaneamente, é influenciado pelo domínio em que está inserido. O leitor não tem a resposta exata do assunto do documento, mesmo porque, a leitura documentária não procura ser, e não deve ser exata. Depende do domínio e de seus objetivos, o que, portanto, relaciona a leitura documentária diretamente à abdução (esta sugere o que pode ser – não há uma certeza, como observado, por exemplo, na dedução).

INFERÊNCIAS	ABDUTIVA	DEDUTIVA	INDUTIVA
DEFINIÇÃO	Criação de hipóteses ou sugestões explicativas sobre os conteúdos do documento	Análise das consequências da atribuição de assunto ao documento	Teste e experimentação com a suposta linguagem do sistema e do usuário
CATEGORIAS	Primeiridade	Secundidade	Terceiridade

NATUREZA	Possibilidade	Generalidade	Continuidade
ETAPAS	Criação	Análise	Comparação
ETAPAS DA LEITURA DOCUMENTÁRIA	Primeiro contato do sujeito com o documento, etapa criativa proporcionada pelo conhecimento de mundo e do domínio. O sujeito já tem uma noção básica do assunto do documento.	Etapa em que ocorre a conferencia da parte criativa por meio da análise da estrutura textual do documento. Consideram-se, também, as particularidades do domínio.	Etapa envolvendo processos finais de Tratamento Temático da Informação (catalogação de assunto e indexação).

Quadro 3: Processo Inferencial de Indexação
Fonte: Adaptado de Almeida, Fujita e Reis (2013)

O conceito de abdução explica de forma consistente o que ocorre durante a leitura documentária, na mente do sujeito. Acredita-se que por meio deste esclarecimento conceitual seja possível fazer uma análise próspera dos processos mentais durante a leitura documentária em um domínio específico. A inferência dedutiva é observada principalmente durante o contato do sujeito com estruturas textuais. A inferência indutiva deve ocorrer após a fase de leitura documentária, na continuidade que ocorre com o contato do usuário com os produtos do tratamento da informação.

Com base nos conceitos apresentados, e em alguns trechos[7] de seis PVIs (nas três áreas do conhecimento), coletados e apresentados em Reis (2012), foi realizada uma análise, relacionando os tipos de inferência dos trechos mais importantes na determinação do assunto principal do documento.

Área de Biológicas 001:

Etapa ABDUTIVA observada na Leitura documentária (trechos):

Então, só de olhar pelo título você já sabe né mais ou menos qual desses cinco assuntos você vai encaixar.

[7] Os PVIs completos podem ser encontrados nos apêndices de Reis (2012)

Etapa DEDUTIVA observada na Leitura documentária (trechos):

Aí nesse aqui eu vou começar pela pagina de rosto, pra leitura técnica que eu faço eu gosto de ler todo o conteúdo {refere-se ao contents do livro}, leio o conteúdo, leio o prefácio, porque por mais que/ por exemplo, o título seja bem explícito, né?

Área de Biológicas 002:

Etapa ABDUTIVA observada na Leitura documentária (trechos):

Agora a próxima obra né, [Ortopedia funcional e mecânica dos maxilares] {leitura do título}. Esse caso aqui já seria um livro mais específico, da área de Ortodontia, e uma experiência a gente já/ seria um livro bem mais específico do que os dois anteriores né, quer dizer, fosse bibliotecário/ começando, a gente acharia que é tudo igual **– Comentário após o primeiro contato com o livro. Comentário sobre o nível de especificidade do livro, e como um bibliotecário com pouca experiência teria dificuldade em realizar a leitura documentária.**

Etapa DEDUTIVA observada na Leitura documentária (trechos):

Agora com relação a esta outra obra né, os mesmos passos né, a gente vai abrir, vai ver a folha de rosto né, como nessa outra também não tem/ o título é [Ortodontia preventiva básica] {leitura do título} do [Alael de Paiva Lima] {leitura do autor do livro} a gente vê a folha de rosto né, daí dá uma olhada no/ faz uma leitura do prefácio

Área de Exatas 001:

Etapa ABDUTIVA observada na Leitura documentária (trechos):

Objeto-orientado, então o livro é básico de geometria algébrica, que é um assunto básico da matemática pura, com aplicações, é um assunto que a gente tem que é objeto-orientado. ***– Comentário após o primeiro contato com o livro, por meio do título.***

Etapa DEDUTIVA observada na Leitura documentária (trechos):

Aí, uma coisa que eu gosto muito de olhar é o prefácio. Porque a parte de prefácio, ela sempre vai me dar informação do livro, por exemplo, se o livro já foi publicado anteriormente, se eles fizeram um estudo de uma coletânea de vários artigos pra basear esse livro

Área de Exatas 002:

Etapa ABDUTIVA observada na Leitura documentária (trechos):

Eu ainda não tô definindo o assunto, mas ele já ta me dando indícios de onde o assunto, o livro vai ficar, que área pelo menos. Eu sei que é da computação. – ***Comentário após o primeiro contato com o livro, por meio do título e do autor.***

Etapa DEDUTIVA observada na Leitura documentária (trechos):

Então esse daqui já tá definido e eu me baseei pela ficha catalográfica que tá no verso da folha de rosto, e também no título, porque sem eu ver a ficha eu já tinha definido.

Área de Humanas 001:

Etapa ABDUTIVA observada na Leitura documentária (trechos):

Ó, esse aqui, [Cogitações sobre o número] {leitura do título}. Do que será isso né? Aí a gente tem que procurar. Aqui, o que que a gente vai ter que fazer, se está falando de cogitações sobre o número, deve ser automaticamente algo sobre o número. Aí o que que eu faço, vou procurando. **– Comentário após o primeiro contato com o livro, por meio do título (inferência abdutiva). Neste caso, a primeira noção do assunto do livro estava errada, o que foi constatado ao buscar informações em outras partes da estrutura textual do livro (inferência dedutiva).**

Etapa DEDUTIVA observada na Leitura documentária (trechos):

[...] tem que olhar aqui dentro mesmo, aqui ó, isso aqui é historia da arte, ó, esse título aqui, esse aqui não tem ficha catalográfica, então o que a gente faz, será que é filosofia? {olha o sumário} Isso aqui vai ser "filosofia" porque fala de dialética, identidade, então a gente vai tirando tudo que tem. **– Verificação do assunto do livro por meio de consulta de estrutura textual.**

Área de Humanas 002:

Etapa ABDUTIVA observada na Leitura documentária (trechos):

[...] então eu vou terminar esse daqui, eu tava no meio, eu tinha visto que era sobre estética na Idade Média, pensando no título [Estética na Idade Média], ele me dá alguma indicação.

Etapa DEDUTIVA observada na Leitura documentária (trechos):

Mas com relação ao livro mesmo eu costumo usar mesmo basicamente a catalogação na fonte, e também no sumário na apresentação.

De acordo com os trechos apresentados acima, é comum observar a inferência abdutiva por meio da consulta do titulo do documento. O leitor proficiente terá uma noção da área do livro ao ler seu titulo (inferência abdutiva), e a consulta às outras partes da estrutura textual ocorrem após esse primeiro contato (inferência dedutiva).

Especificamente na área de Humanas, observa-se com mais frequência a inferência abdutiva, mesmo que incorreta, tornando a inferência dedutiva essencial no processo de leitura documentária. Ambas se completam, uma leitura que foca em apenas um tipo de inferência, deve ser considerada incompleta.

Os quadros 2, 3 e a análise dos trechos de leitura apresentados, são evidencias de que é possível categorizar o pensamento verbalizado de sujeitos durante a leitura documentária, utilizando metodologias especificas e a teoria da semiótica. Tais constatações foram proporcionadas por

resultados de pesquisas anteriores e estudos teóricos apresentados aqui. Recomenda-se a aplicação em um domínio especifico da área de Humanas.

Considerações finais

Com base nos conceitos expostos sobre os aspectos semióticos da leitura documentária, pode-se concluir que de fato, a leitura documentária, é um processo inferencial, e que ocorre por meio da abdução e da dedução. As etapas finais dos processos de tratamento temático da informação, somente são possíveis graças ao uso de estratégias inferenciais indutivas por parte dos leitores proficientes (bibliotecários).

Recomenda-se a aplicação das teorias semióticas em domínios específicos (principalmente da área de Humanas – pois, conforme observado por Reis (2012), a leitura documentária no tratamento temático é mais complexa nesta área). Acredita-se que por meio de análises baseadas em teorias da semiótica, seja possível mapear de forma mais adequada – aos produtos finais do tratamento temático da informação – os processos metacognitivos em leitura documentária de cada profissional.

A intenção não é criar um modelo ou padrão, e sim, realizar uma tentativa de esquematizar como leitores proficientes realizam a leitura documentária em determinados contextos com base em livros específicos de uma área e no comportamento do sujeito durante a atividade. Em pesquisas anteriores, o livro (suporte tradicional), foi o material usado para análise, sem o cuidado de observar que tipos de livros existem dentro de cada área (Ex: Humanas: literatura ou conteúdo? Em Direito? Em Psicologia? Qual área específica?).

Sugere-se que a aplicação seja detalhada, observando os tipos de livros a serem tratados, com base em teorias da semiótica, especificamente, a abdução e dedução.

Referencias

ALMEIDA, C. C. Conceito como signo: elemento semiótico para análise e mediação da informação. *Scire*, Saragoza, v. 2, n. 18, p.49-55, jun. 2012.

ALMEIDA, C. C.; FUJITA, M. S. L.; REIS, D. M. Peircean Semiotics and Subject Indexing: Contributions of Speculative Grammar and Pure Logic. *Knowl. Org.*, v. 40, n. 4, p.225-241, 2013.

CINTRA, A. M. M. Elementos de lingüística para estudos de indexação. *Ciência da Informação*, v. 12, n. 1, p. 5-22, 1983.

DIAS, E. W.; NAVES, M. M. L. *Análise de assunto*: teoria e prática. 2. ed. Brasília: Briquet de Lemos, 2013.

DIAS, E. W.; NEVES, D. A. B.; PINHEIRO, A. M. V. Uso de estratégias metacognitivas na leitura do indexador. *Ciência da Informação*, Brasília, v. 35, n. 3, p. 141-152, set./dez. 2006.

ECO, U. *Semiótica e filosofia da linguagem*. São Paulo: Piaget, 2001. (Teoria das artes e literatura).

ECO, U. *Tratado geral de semiótica*. 5. ed. São Paulo: Perspectiva, 2014. (Estudos).

ELLIS, A. W. *Leitura, escrita e dislexia*: uma análise cognitiva. 2. ed. Porto Alegre: Artes Médicas, 1995.

FUJITA, M. S. L. A leitura documentária na perspectiva de suas variáveis: leitor-texto-contexto. *Datagramazero*, Rio de Janeiro, v. 5, n. 4, 2004.

FUJITA, M. S. L.; NARDI, M. I. A.; FAGUNDES, S. A. A observação da leitura documentária por meio de protocolo verbal. In: RODRIGUES, G. M., LOPES, I. L. (Org.) *Organização e representação do conhecimento na perspectiva da ciência da Informação.* Brasília: Thesaurus, 2003. p. 141-178. (Estudos Avançados em Ciência da Informação, v.2).

FUJITA, M. S. L.; NARDI, M. I. A.; SANTOS, S. A. S. A leitura em análise documentária. *Transinformação*, v. 10, n. 3, p. 13-31, 1998. Disponível em: <http://www.brapci.ufpr.br/brapci/v/a/285>. Acesso em: 11 Jan. 2017.

HJØRLAND, B. Fundamentals of knowledge organization. *Know. Org.*, v. 30, n. 2, p. 87-111, 2003.

KATO, M. A. *No mundo da escrita*: uma perspectiva psicolinguística. 5. ed. São Paulo: Ática, 1995. (Fundamentos).

KOCH, I. V. *O texto e a construção dos sentidos*. 2. ed. São Paulo: Contexto, 1998.

LIVINGSTON, J. A. A. *Metacognition*: an overview. Copyright 1997. Disponível em: <http://www.gse.buffalo.edu/fas/shuell/cep564/Metacog.htm>. Acesso em: 11 Jan. 2017

MAI, J-E. Semiotics and indexing: na analysis of the subject indexing process. *Journal of Documentation*, v. 57, n. 5, p. 591-622, 2001.

MAI, J-E. Analysis in indexing: document and domain centered approaches. *Information Processing And Management*: An International Journal, Nova York, v. 41, p. 599-611, fev. 2004.

NERIS, L. O. Semiótica e leitura: o fazer-receptivo do leitor analista. *Estudos Semióticos*, Número 2, São Paulo, 2006. Disponível em <www.fflch.usp.br/dl/semiotica/es>. Acesso em "14/12/2016"

NEVES, D. A. B. Leitura e metacognição: uma experiência em sala de aula. *Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf.*, Florianópolis, n. 24, p. 1-9, 2º sem. 2007.

PEIRCE, C. S. *Collected Papers of Charles Sanders Peirce*. Ed. Hartshorne, Charles; Weiss, Paul; Burks, Arthur. Cambridge, MA: Harvard University Press, 1931-1958. 8v.

PIGNATARI, D. *Semiótica e literatura*. 6. ed. Cotia: Ateliê Editorial, 2004.

PINTO, M.; GÁLVEZ, C. *Análisis documental de contenido*. Madrid: Síntesis, 1999.

REIS, D. M. *A importância da observação da estrutura textual durante a catalogação de assunto de livros científicos em bibliotecas universitárias*: uma análise realizada a partir da técnica de Protocolo Verbal. 2012. 132 f. Dissertação (Mestrado) - Curso de Mestrado em Ciência da Informação, Programa de Pós-graduação em Ciência da Informação da Faculdade de Filosofia e Ciências da Unesp, Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Marília, 2012. Disponível em: <http://repositorio.unesp.br/handle/11449/93675>. Acesso em: 11 abr. 2017.

THELLEFSEN, T. Semiotic Knowledge Organization: theory and method development. *Semiotica*, v. 142, p. 71-90, 2002.

APÊNDICE 1

MODELO DE LEITURA PARA INDEXAÇÃO DE TEXTOS CIENTÍFICOS: MANUAL EXPLICATIVO[1]

O manual explicativo contém o modelo de leitura acompanhado de explicações, a cada etapa de sua realização, seguido de muitos exemplos que ilustram sua aplicabilidade. De acordo com o Manual, a orientação para a leitura está dividida em três etapas: exploração da estrutura textual, identificação de conceitos e seleção de conceitos. Em cada uma destas etapas os procedimentos são esclarecidos um após o outro, deixando transparecer uma preparação conceitual e filosófica sobre indexação baseada nos resultados da pesquisa.

Apresenta-se, a seguir, o manual explicativo do "Modelo de leitura para indexação, cuja demonstração é feita a partir de orientações explicativas para que indexadores possam realizar um "passo-a-passo" da tarefa.

MODELO DE LEITURA PARA INDEXAÇÃO DE TEXTOS CIENTÍFICOS

Manual explicativo

---

[1] Sintetizado e adaptado da tese de Livre-Docência: (FUJITA, 2003, p. 226-243);
Parte do Capítulo "A leitura documentária e o processo de compreensão do indexador: memorial de investigação científica".
Publicado em:
FUJITA, M. S. L., RUBI, M. P. Um modelo de leitura documentária para a indexação de artigos científicos: princípios de elaboração e uso para a formação de indexadores. *Data Grama Zero***,** Rio de Janeiro, v. 7, n.3, p. 1-18, 2006a.
FUJITA, M. S. L., RUBI, M. P. Modelo de lectura profesional para la indización. *Scire* (Zaragoza), v.12, p. 47 – 70, 2006b.

A leitura documentária, realizada pelo indexador na fase de análise, visa a propiciar a "identificação de conceitos" para posterior representação em índices que satisfaçam a demanda do usuário.

A indexação em análise documentária, sob o ponto de vista dos sistemas de recuperação de informação, é reconhecida como a parte mais importante porque condiciona os resultados de uma estratégia de busca. *O bom ou mau desempenho da indexação reflete-se na recuperação da informação feita por meio de índices.* Isso nos leva a considerar que a recuperação do documento mais pertinente à questão da busca é aquele cuja indexação proporcionou a identificação de conceitos mais pertinentes ao seu conteúdo, produzindo uma correspondência precisa com o assunto pesquisado em índices.

Na identificação de conceitos, o indexador, após o exame do texto, passa a abordá-lo de uma forma mais lógica a fim de selecionar os conceitos que melhor representem seu conteúdo. E a seleção de termos é necessária tendo em vista os objetivos para os quais as informações são indexadas. Assim, nem todos os conceitos identificados serão necessariamente selecionados.

No contexto da análise para indexação, a leitura constitui a atividade principal da indexação, pois é a fase inicial que influenciará no desempenho de outras operações e resultará na seleção de termos que irão representar o documento para o usuário. Assim, a leitura passa a ser mais direcionada aos objetivos de indexação, diferente, pois, da leitura para outros fins.

Considerando-se que a identificação de conceitos é o objetivo da leitura documentária e a operação mais importante da indexação, concluímos que:

- a identificação de conceitos pode depender do domínio do indexador na exploração da estrutura textual;
- existem duas operações distintas realizadas pelos indexadores ***durante*** (e não após!) a leitura – Identificação de conceitos e Seleção de termos.

Assim, essa instrução de leitura estará dividida em três procedimentos principais:

I. Exploração do conhecimento da estrutura textual

II. Identificação de conceitos

III. Seleção de conceitos

EXPLORAÇÃO DO CONHECIMENTO DE ESTRUTURA TEXTUAL

1. Observação da estrutura textual: ize no texto científico os elementos que o compõem, mesmo que não estejam evidentes nos itens ou sinalizados por meio de destaques. Todo texto possui uma estrutura, evidente ou não, que podemos denominar de superestrutura. Essa observação deve ser feita, portanto, com base na superestrutura, pois indicará, com maior objetividade, qual é o assunto tratado no texto, evitando que se cometam equívocos:

- Título em português
- Título em inglês
- Autoria
- Resumo do trabalho científico
- Palavras-chave
- Abstract
- Keywords
- Introdução
- Materiais e métodos
- Resultados
- Figuras
- Discussão dos resultados
- Conclusões
- Referências bibliográficas

1. Localização do conteúdo pertinente de cada uma dessas partes do texto. Verifique que o conteúdo pertinente a cada parte do texto demonstra um padrão, tal como:

   - Introdução: explicação do assunto principal com referencial teórico, contendo os objetivos com o tema principal do trabalho ao final da introdução;

   - Materiais e métodos: descrição de materiais e métodos utilizados, processos, técnicas, amostragem;

   - Resultados: compatibilidade com objetivos enunciados e materiais e métodos utilizados, utilizando , às vezes, figuras, gráficos, tabelas, fotografias etc.;

   - Discussão dos resultados: verificação dos resultados a partir do referencial teórico utilizado;

   - Conclusões: verificação dos objetivos propostos;

   - Referências bibliográficas

IMPORTANTE: *a realização da **etapa 2** é **imprescindível**, pois resultará na compreensão global do texto.*

II. Identificação de conceitos

A metodologia utilizada para identificação de conceitos será realizada combinando a exploração da estrutura textual e o questionamento.

A identificação de conceitos é a etapa principal da indexação e dependerá da compreensão do que é conceito e qual a sua importância. Conceito é a formulação de uma idéia por palavras. Tomemos como exemplo o conceito *agente* que pode ser definido por *aquele ou algo que realizou a ação*. Isso significa que o conceito agente poderá ser representado por uma palavra no texto, que dependerá do contexto, para identificar-se com a idéia de agente. Assim, asseguramos que esses conceitos poderão ser identificados em qualquer texto, garantindo uma uniformidade de identificação de conceitos e de compreensão global do texto que, de outra forma, não seria possível por não termos parâmetros de compreensão.

Ex.: Destruição de plantações de café pela geada.

O agente neste caso é a geada, pois ela "praticou" a ação de destruição de plantações de café.

↳ *Compreensão de conceitos*

Os conceitos essenciais do documento são:

- OBJETO: é algo ou alguém que está sob estudo do pesquisador.
- AÇÃO: processo sofrido por algo ou alguém
- AGENTE: aquele ou algo que realizou a ação
- MÉTODOS: métodos utilizados para realização da pesquisa
- FÍSICO OU AMBIÊNCIA: físico onde foi realizada a pesquisa
- CAUSA E EFEITO:

- *causa* => razão ou motivo. Aquilo ou aquele que faz com que uma coisa exista ou aconteça (antecedente); está vinculada à identificação da AÇÃO.

- *efeito* => produto de uma causa. Resultado de um ato qualquer (conseqüente); está vinculado ao resultado da AÇÃO realizada.

Assim, o suposto efeito ou conseqüente deve variar cada vez que varia a suposta causa ou antecedente.

Ex.: Grupos familiares carentes que migram tendem a desorganização interna.

Causa: processo de migração

Efeito: desorganização interna do grupo familiar

↳ *Identificação de conceitos mediante exploração da estrutura textual*

Para conseguir um melhor resultado na identificação de conceitos, você poderá utilizar partes do texto em que os conceitos, geralmente, poderão ser identificados:

PARTES DO TEXTO	CONTEÚDO PERTINENTE	CONCEITOS DE
Introdução (objetivos)	Referencial teórico Tema: objetivos* Objetivos	OBJETO AGENTE AÇÃO
Metodologia	Descrição de materiais, métodos, processos e técnicas utilizados.	MÉTODOS FÍSICO MATERIAIS
Resultados **Discussão dos resultados**	Compatibilidade com objetivos enunciados e materiais e métodos utilizados, mostrados, às vezes em tabelas; Verificação dos resultados a partir do referencial teórico utilizado	CAUSA E EFEITO

Quadro 1: **Identificação de conceitos mediante exploração da estrutura textual**
* observe que o tema, geralmente, está expresso no objetivo.

↳ *Questionamento do texto para identificação*

Por outro lado, esse resultado poderá ser obtido mais facilmente se você utilizar o questionamento a seguir, pois as respostas a essas perguntas implicarão uma análise do documento e darão origem à seleção de termos.

Abaixo temos um exemplo que demonstra o uso do questionamento e a obtenção de termos como resposta à identificação dos conceitos estabelecidos.

Ex.: Proliferação da flora anaeróbia no intestino delgado em lactentes portadores de diarréia aguda e persistente.

1. O ASSUNTO CONTÉM UMA AÇÃO (PODENDO SIGNIFICAR UMA OPERAÇÃO, UM PROCESSO ETC.)?

=> AÇÃO: proliferação

2. O DOCUMENTO POSSUI EM SEU CONTEXTO UM OBJETO SOB EFEITO DESTA AÇÃO?

=> OBJETO: flora anaeróbia

2.1 O OBJETO IDENTIFICADO PODE SER CONSIDERADO COMO PARTE DE UMA TOTALIDADE?

=> PARTE DO OBJETO: "flora anaeróbia" é parte do "intestino delgado" que é parte do todo "lactente"

2.2 O OBJETO IDENTIFICADO POSSUI CARACTERÍSTICAS OU ATRIBUTOS PARTICULARES?

No exemplo dado não existe característica ou atributo, mas em outro exemplo, seria:

Substância aromática do vinho

vinho: objeto

substância aromática: atributo

3. O DOCUMENTO POSSUI UM AGENTE QUE PRATICOU ESTA AÇÃO?

=> AGENTE: microorganismos anaeróbios

4. PARA ESTUDO DO OBJETO OU IMPLEMENTAÇÃO DA AÇÃO, O DOCUMENTO CITA E/OU DESCREVE MODOS ESPECÍFICOS, POR EXEMPLO: INSTRUMENTOS ESPECIAIS, TÉCNICAS, MÉTODOS. MATERIAIS E EQUIPAMENTOS?

=> MÉTODOS: Intubação intestinal; análise morfológica das colônias

=> MATERIAIS:

5. A AÇÃO, OBJETO E AGENTE SÃO CONSIDERADOS NO CONTEXTO DE UM LUGAR ESPECÍFICO OU AMBIENTE?

   => FÍSICO OU AMBIÊNCIA: a pesquisa foi realizada em Unidades de Gastroenterologia Pediátrica

6. CONSIDERANDO QUE A AÇÃO E O OBJETO IDENTIFICAM UMA CAUSA, QUAL É O EFEITO DESTA CAUSA?

   => CAUSA: proliferação da flora anaeróbia (AÇÃO+OBJETO);

   => EFEITO: diarréia aguda e persistente, pois quando há aumento da proliferação da flora anaeróbia, agrava-se diarréia aguda e persistente.

Observação: às vezes, nem todas as questões poderão ser respondidas.

Após a compreensão do texto citado, faça a identificação de conceitos combinando a exploração da estrutura textual com o questionamento.

III. Seleção de conceitos

A partir da identificação de conceitos, realizada por meio das respostas a essas questões, selecione os conceitos que você considera importantes para uma representação mais pertinente ao conteúdo do documento e que constem do sistema, conforme linguagem adotada, promovendo a garantia de uso do documento.

Exemplo:

TERMOS IDENTIFICADOS	TERMOS SELECIONADOS
Flora anaeróbia	Flora anaeróbia
Proliferação	Proliferação da flora anaeróbia
Microorganismos anaeróbios	Microorganismos anaeróbios
Intubação intestinal	Análise morfológica
Análise morfológica das colônias	Diarréia aguda
Unidades de Gastroenterologia Pediátrica	Diarréia persistente
Proliferação da flora anaeróbia	
Diarréia aguda e persistente	

Quadro 2: Identificação e seleção de termos

APÊNDICE 2

DIRETRIZES TEÓRICO-METODOLÓGICAS SOBRE LEITURA DOCUMENTÁRIA PARA INDEXAÇÃO[2]

Mariângela Spotti Lopes Fujita

O Programa de orientação à formação do indexador em leitura documentária contém os principais subsídios teóricos e metodológicos das investigações dos Eixos I e II que fundamentam o aprimoramento do ensino de indexação.

Os subsídios mais importantes coletados referem-se aos seguintes itens: Leitura e estratégias de leitura; O texto e sua estrutura; O indexador como leitor: suas estratégias e conhecimento prévio; O leitor profissional e seu contexto. A organização desses subsídios pautou-se pela perspectiva interacionista das variáveis que participam do processo de leitura: o texto, o leitor e o contexto, produzindo a estrutura básica: Leitura em análise documentária; As variáveis influentes na leitura: o texto, o indexador como leitor e o contexto.

[2] Sintetizado e adaptado da tese de Livre-Docência: (FUJITA, 2003, p. 244-251): Parte do Capítulo "A leitura documentária e o processo de compreensão do indexador: memorial de investigação científica".

1 Leitura e estratégias de leitura

A leitura, segundo Kleiman (2000), é um *ato individual* realizado por apenas um leitor, porém, é também um *ato social* porque existe um processo de comunicação entre o leitor e o autor do texto, ambos com objetivos estabelecidos anteriormente dentro do contexto de cada um.

O leitor traz consigo seu conhecimento prévio, suas experiências acumuladas, seus valores, e utiliza essa bagagem para interagir com o texto (os pontos de vista, as intenções do autor e suas idéias implícitas no texto) (CAVALCANTI, 1989).

Com visão semelhante à de Cavalcanti, Giasson (1993) acredita que o leitor cria sentido, apoiando-se simultaneamente no texto, nos seus conhecimentos prévios e na intenção da leitura. Qualquer processo de compreensão de texto escrito, portanto, é um ato de comunicação que envolve essas três variáveis: o leitor, o texto contendo as idéias do autor e o contexto e variará de acordo com o grau de relação entre elas.

A interação entre as três variáveis desenvolve-se durante o processo de compreensão de leitura, principalmente, pelo uso de estratégias entendidas como ações conscientes do leitor (metacognitivas) direcionadas para um objetivo ou para busca de solução de problemas de compreensão ou como ações subconscientes (cognitivas) durante a leitura fluida. Durante a leitura são ativados esquemas variados, desde conhecimento de vocabulário, conhecimento da estrutura textual, do assunto, até conhecimento de mundo. O conhecimento prévio, importante para a compreensão, é aquele existente na memória a longo prazo, a qual é permanente, ilimitada e constituída de esquemas (representações generalizadas), que podem ser ativados durante o processo de leitura por dois movimentos complementares interativos: *bottom-up* e *top-down*.

A característica principal deste modelo do processo de leitura de Rumelhart (1977) é a interação dos componentes de cada nível, a ativação de todos os subprocessos e a influência da informação contextual em todo o processamento. Pinto e Gálvez ([1996], p. 45) consideram o modelo interativo de Rumelhart (1977) o que melhor representa a atuação do sujeito documentalista porque:

> ... quando está lendo um texto cujo assunto lhe seja desconhecido, procederá dos níveis mais inferiores ao superiores, fará uma leitura lenta, detalhada, ascendente ou bottom-up. Ao contrário, quando o texto for familiar, realizará uma leitura "entre linhas" - descendente ou top-down -, antecipando informações e dirigindo-se a uma representação do conteúdo global do texto.

A construção do sentido do texto, segundo Kleiman (1989, p. 13), depende da "[...] interação de diversos níveis de conhecimento, como o conhecimento lingüístico, o textual e o conhecimento de mundo." Segundo a autora, o conhecimento textual desempenha um papel importante na compreensão de textos, pois o domínio da tipologia documentária e da estrutura textual são dois tipos de conhecimento prévio que poderão aumentar compreensão do indexador durante o processo descendente de leitura.

O mais importante, porém, é que o leitor possa monitorar sua compreensão durante a leitura e isso é possível pela metacognição, envolvendo um conjunto de estratégias metacognitivas. A metacognição em leitura permite ao leitor uma compreensão de sua própria compreensão, ou melhor, que ele possa acompanhar e avaliar seu processo de compreensão durante a leitura de um texto e, além disso, tomar providências quando a compreensão falha (LEFFA, 1996, p.45). Além das estratégias metacognitivas existem as cognitivas que são características em uma leitura mais fluida. O uso de estratégias cognitivas e metacognitivas será alternado durante a leitura e sempre tenderá a um equilíbrio.

O uso de estratégias ainda que não seja facilmente observável, porque ações mentais, como associações e deduções durante a leitura, não podem ser vistas, pode ser verbalizado. Para observar essas estratégias durante a leitura, existem métodos introspectivos, um dos quais é o Protocolo Verbal, inicialmente utilizado em Psicologia Cognitiva e depois em Lingüística Aplicada. Para conferir natureza metacognitiva às ações mentais, Brown (1980, p.456) indica as seguintes atividades:

> explicitação dos objetivos da leitura; identificação de aspectos importantes da mensagem; alocamento de atenção a áreas importantes; monitoração do comportamento para ver se está ocorrendo compreensão; engajamento em revisão e auto-indagação para ver se o objetivo está

> sendo atingido; tomada de ações corretivas quando são detectadas falhas na compreensão; recobramento de atenção quando a mente se distrai ou faz digressões.

O estudo sobre "Leitura" e "Estratégias de Leitura" revelou que a leitura para fins documentários exige que o leitor seja metacognitivo. O leitor documentalista deve fazer uso principalmente de estratégias metacognitivas tais como a exploração do seu conhecimento de estruturas textuais e deve manter em mente o seu objetivo de representar o texto para futura recuperação, considerando as limitações da tarefa de indexação e os objetivos do sistema de informação no qual se insere.

Considerando-se que a leitura documentária desenvolve-se durante a análise de assunto, certamente, a formação do indexador deve incluir a compreensão do processo de análise, visto que é etapa mais importante porque dela deverá resultar a representação do assunto principal do documento.

Conforme resultados da observação de estratégias de leitura, está explícito que existem duas operações distintas utilizadas pelos indexadores durante a leitura: *Identificação de conceitos* e *Seleção de conceitos.* A seleção de conceitos ocorre em dois momentos diferentes da análise de assunto: durante a identificação de conceitos para a determinação do assunto e, após a identificação de conceitos, durante a tradução dos termos que representam os conceitos para os termos da linguagem documentária adotada pelo serviço de análise.

Destaca-se, entretanto, que a análise de assunto reveste-se de uma subjetividade característica, dadas as circunstâncias e elementos envolvidos, pois é realizada a partir da **leitura** do documento pelo indexador. A questão da identificação do assunto do documento está, pois, vinculada à leitura. A **leitura** é uma atividade autônoma que instrumentaliza a análise de assunto. No momento em que o indexador busca efetivamente o assunto, começa a intersecção com a análise. Antes dessa busca, o leitor observa a estrutura textual para descobrir em quais partes do texto encontrará os conceitos a serem identificados e selecionados durante a análise de assunto.

A análise de assunto implica a determinação da tematicidade do documento mediante a identificação e seleção dos conceitos que comporão o assunto ou tema principal e secundários.

O tema possui uma estrutura composta por conceitos ou categorias ou facetas. A identificação das categorias ou conceitos na estrutura textual do documento decorrerá da análise conceitual e a composição das categorias identificadas formulará o tema do documento.

A identificação da estrutura temática, composta pelos conceitos, leva à identificação do tema. Contudo, dependendo da legibilidade e da estrutura textual, o tema poderá estar formulado no objetivo do trabalho de forma clara ou, quando isto não acontecer, a identificação dos conceitos deverá ser feita dentro da estrutura textual do documento.

Portanto, podemos considerar que a primeira etapa do processo de indexação, a análise de assunto, constitui-se das seguintes fases:

- Determinação da tematicidade intrínseca
- Identificação de Conceitos
- Seleção de Conceitos
- Determinação da tematicidade extrínseca

Conforme Albrechtsen (1993, p.221), dependendo dos objetivos institucionais, percebe-se qual a concepção de análise de assunto que o sistema de informação segue e, conseqüentemente, o indexador levará esse aspecto em questão. Consideram-se, assim, diferentes concepções de análise que, certamente afetam o desempenho do indexador enquanto leitor. A esse respeito, Albrechtsen (1993, p.220) considera três diferentes concepções de análise de assunto: a concepção simplista, a orientada para o conteúdo e a orientada para a demanda.

A concepção simplista lida com as informações explícitas dos documentos, considerando o assunto como abstração direta dos documentos.

A concepção orientada para o conteúdo envolve uma interpretação do conteúdo do documento que vai além da estrutura léxica e, às vezes, gramatical do texto, ou seja, envolve aspectos mais complexos que o processo da concepção simplista.

A concepção orientada para a demanda envolve os dados do assunto como instrumentos de transferência do conhecimento, apontando um encontro pragmático da informação ou do conhecimento. Este tipo de indexação envolve um alto grau de responsabilidade por parte do indexador, ao julgar a qualidade do documento para usuários potenciais, tornando-se necessário antecipar a demanda, indo além das fronteiras que separam acervo e usuários.

De acordo com Naves (1996), as duas últimas concepções - orientada pelo conteúdo e orientada pela demanda – são complementares. Consideramos que são mais do que complementares, são intrínsecas porque no momento em que o indexador está lendo e procurando identificar e selecionar conceitos para a determinação do assunto do documento está objetivando encontrar o assunto que lhe é familiar devido à sua prática de indexação e também o que pode interessar ao usuário do sistema de informação.

A concepção de leitura orientada para o conteúdo deve orientar a identificação de conceitos e a concepção orientada para a demanda, a seleção de conceitos durante a leitura documentária.

2 As variáveis influentes na leitura: o texto, o leitor e o contexto

Na variável *texto*, constatou-se que o conhecimento da estrutura textual permite ao indexador uma estratégia que facilita a leitura e compreensão do conteúdo e agilidade na leitura, pois em muitos momentos da leitura o indexador pode praticamente "saltar" de um trecho a outro para buscar o que precisa. Pela variável *leitor,* considerou-se que o indexador torna-se um leitor no ato de análise de um documento com a finalidade de realizar a indexação para representação do conteúdo por termos que serão, posteriormente, recuperados por um usuário do sistema de informação.

Na variável *contexto,* Giasson (1993) indicou três elementos extratexto: *contexto psicológico* (intenção de leitura, interesse pelo texto), *contexto social* (intervenção dos colegas e dos chefes imediatos) e o *contexto físico* (o tempo disponível, as condições materiais).

2.1 O texto

Dentre as várias acepções do que vem a ser o texto, Koch (2002, p. 17) o coloca na concepção interacional da língua, e considera o texto como o lugar da interação quando se tem, junto ao sujeito leitor, o contexto sociocognitivo dos participantes da interação. Nesta concepção, Koch (2002, p. 20) compartilha e subscreve a definição proposta por Beaugrande (1997, p.10) para texto: "evento comunicativo no qual convergem ações lingüísticas, cognitivas e sociais."

> No que diz respeito à estrutura do texto, afirma-se estar associado ao modo com o qual as idéias são organizadas no texto; com relação ao conteúdo, diz-se estar associado ao tema, aos conceitos tratados no texto. A estrutura do texto se articula ao seu conteúdo, para tanto, o autor de um texto escolherá determinada estrutura textual que venha coincidir com o conteúdo que quer transmitir.

A idéia principal varia de acordo com a estrutura textual, por exemplo: num texto narrativo, a idéia principal pode ser um acontecimento ou a sua interpretação; num texto informativo, pode ser uma regra, um conceito, ou uma generalização. Quando a idéia principal aparece implícita, o leitor deve inferi-la com base nas informações fornecidas pelo texto e no seu conhecimento prévio sobre o assunto.

Assim, além de uma estrutura textual e lingüística, o texto possui uma estrutura de significado que só pode ser identificada, ou que somente "aparece", quando o leitor faz uma leitura compreensiva. Essa ligação de estrutura ao conteúdo, "perceptível" durante a leitura, indica que todo texto possui uma superestrutura e uma macroestrutura que o caracteriza. Van Dijk (1992, p.142) explica metaforicamente que uma superestrutura é um tipo *de forma* do texto, cujo objeto, o tema, isto é, a macroestrutura, é o conteúdo do texto.

Esse tipo de conhecimento prévio pelo leitor (de estruturas textuais) possibilita-lhe a identificar a parte do texto que traz a idéia principal auxiliando-o, assim, a compreender, de forma global, o texto e a realizar uma leitura, pois já conhece as partes que tem a explorar e os conceitos pertencentes a cada parte, chegando, dessa forma, ao tema do texto.

Observando a variável texto no processo de leitura, numa perspectiva macro, é possível notar que os textos apresentam uma estrutura com as partes informacionais que os compõem organizadas numa seqüência lógica diferenciada de uma tipologia textual para outra.

Ao abordar a estrutura dos textos, Van Dijk (1992) enfatiza as estruturas narrativa, argumentativa e a do discurso científico, considerando que os textos narrativos são formas básicas muito importantes da comunicação textual; as argumentativas são as mais utilizadas em filosofia e teoria da lógica e servem de base ao discurso científico.

Considerando que este programa privilegia, em um primeiro momento, orientações para leitura do texto científico, passaremos a abordá-lo mais objetivamente. Kobashi (1994, p.114-6) apresenta a seguinte organização básica para o texto cientifico:

> **Tema**: assunto que se irá desenvolver
>
> **Problema**s: dificuldade que se quer solucionar cientificamente
>
> **Hipótese**: proposição que se antecipa à comprovação de uma realidade
>
> **Metodologia**: procedimentos e operações que possibilitem a observação racional e controlada dos fatos, de modo a permitir a interpretação e a explicação adequada do fenômeno observado.
>
> **Resultado**: implica a aceitação ou não das hipóteses formuladas ou na reformulação das mesmas;
>
> **Conclusão**: comentário final, que se discute as possibilidades de aplicação e de utilização dos resultados, isto é, a incorporação ou não destes últimos a um sistema teórico.

Segundo resultados e conclusões (FUJITA, 1998, 1999), obtidos a partir de observações dos Protocolos verbais, o domínio da estrutura textual pelo indexador faz parte de seu conhecimento prévio textual porque

foi identificado como estratégia de leitura que facilita a tarefa de identificação e seleção de conceitos durante a análise de assunto.

As partes da estrutura textual mais exploradas pelos sujeitos indexadores (num total de oito sujeitos) e em que foi possível encontrar os conceitos pertencentes ao tema, foram, em ordem seqüencial da estrutura: título, resumo, introdução, material e métodos, resultados, conclusão e referências bibliográficas. Por se tratar de um estudo realizado em áreas de assunto em que os títulos são reais, não se encontram títulos metafóricos. Essa parte da estrutura foi, juntamente com o resumo, as mais utilizadas para identificação e seleção dos conceitos representativos do tema do texto. O título foi o mais explorado, seguido do resumo e da introdução, depois a conclusão, material e métodos, referências bibliográficas e outras partes do texto como: tabelas, gráficos etc.

Consideramos, portanto, que é necessária a elaboração de um modelo de leitura que utilize o conhecimento de estruturas textuais do leitor apoiado em estratégias mais sistemáticas de identificação de conceitos, para efetuar não só o movimento ascendente (bottom-up), mas a interação entre os dois movimentos de leitura - ascendente e descendente ao mesmo tempo, preconizado por Rumelhart (1977) e considerado o que melhor representa a atuação do sujeito documentalista, de acordo com Pinto e Gálvez ([1996]).

2.2 O leitor

Entre as variáveis influentes no processo de leitura, o Indexador (como leitor) deve ser considerado a variável mais importante, pois, além de um leitor inato, é um leitor com objetivo profissional, que possui uma concepção de análise de assunto baseada em sua formação na graduação e nos cursos de capacitação. Pela análise sobre sua formação, verificou-se que o indexador foi formado e capacitado para uma análise de texto mais orientada pelas linguagens documentárias do que para o conteúdo. Tratase de um leitor que interage com o texto e, a identificação de conceitos depende de estratégias que facilitem sua compreensão.

Durante a leitura, o leitor-indexador faz uso de seu conhecimento prévio implícito: (lingüístico, textual e de mundo), utiliza estratégias meta-

cognitivas (que permitem uma compreensão de sua própria compreensão), assim como estratégias cognitivas (conhecimentos da ordem natural, sintática e semântica), depositando sobre o texto expectativas acerca da coerência textual de forma global e . Sendo, em princípio, o indexador um leitor apto à compreensão, uma maior compreensão de leitura por parte dele implica no desenvolvimento de seu auto-conhecimento em leitura, habilidades e estratégias de leitura e aprofundamento de seu conhecimento prévio.

O estudo sobre observação da leitura do indexador identificou estratégias utilizadas e a intervenção do conhecimento prévio no processo de identificação e seleção de conceitos. Os resultados revelaram que a leitura dos indexadores é metacognitiva ao explorar toda a estrutura textual do documento, ao utilizar estratégias de leitura, principalmente, ao manter em mente o objetivo da leitura – representar para recuperar.

A principal constatação a respeito da observação de estratégias de leitura foi a de que existem duas operações distintas utilizadas pelos indexadores durante a leitura: Identificação de conceitos e Seleção de conceitos. Isso esclarece que o indexador realiza as duas operações durante a leitura e não após a mesma. Para diminuir as dificuldades durante a identificação de conceitos é necessário realizar uma leitura compreensiva dotada de estratégias que permitam agilizar e facilitar a identificação de conceitos.

Verificou-se, na investigação, que o indexador possui objetivos profissionais claros quanto à análise para condensação documentária: representar o conteúdo documentário, com palavras-chaves, para futura recuperação pelo usuário do sistema de informação. Por conta de seus objetivos deve ser entendido como **leitor profissional,** dotado de conhecimento prévio sobre tipologias documentárias e estruturas textuais, estratégias metacognitivas de leitura documentária voltadas para seus objetivos e conhecimento prévio específico da linguagem documentária do sistema para domínio de áreas do conhecimento especializado.

Isso significa que, mesmo sem conhecimento prévio específico, ele poderá desenvolver a compreensão do texto e realizar a análise de forma conceitual, caso ele desenvolva habilidades e estratégias de leitura documentária, aprofunde seu conhecimento prévio lingüístico e textual, desenvolva uma experiência de análise em área especializada, com domínio da linguagem do-

cumentária, para obter familiaridade com o assunto, e principalmente, tenha uma formação em indexação capaz de incutir-lhe a necessidade e a importância de realizar uma análise conceitual pertinente e, ainda, seja consciente de suas habilidades cognitivas inatas e construídas. Torna-se necessário, portanto, ao indexador, compreender o processo de análise documentária, definir os objetivos para a leitura documentária e dominar a linguagem documentária para conhecimento prévio da área de assunto.

A concepção de análise assumida pelo indexador ao realizar a análise de um texto está diretamente vinculada com a formação educacional e com a política do sistema de informação e não pelo fato de ele ser um leitor menos ou mais habilitado.

Na formação do indexador brasileiro, constatou-se que a atividade está mais articulada com o desenvolvimento da prática e menos com a fundamentação teórica e que os conteúdos programáticos de disciplinas da matéria "Tratamento Temático da Informação" são mais dedicados ao uso de linguagens documentárias, preocupando-se cada vez mais com a formação do indexador em leitura documentária, o que, certamente, tenderá a desenvolver o tema "leitura documentária" em futuras reformulações curriculares. É necessária a inclusão de conteúdo sobre leitura documentária, adotando-se a visão interacionista do processo de leitura através das três variáveis: o texto, o leitor e o contexto.

Além disso, é preciso adotar bibliografia específica sobre leitura e preocupar-se com a geração de conhecimento teórico e metodológico. É importante demonstrar a realidade de atuação profissional do indexador em sistemas de informação e conhecer as áreas especializadas e suas linguagens documentárias.

Com base nas três concepções de análise de assunto (ALBRECHTSEN, 1993) que afetam o desempenho do indexador enquanto leitor (simplista, orientada para o conteúdo e orientada para a demanda) realizou-se descrição das estratégias de cada concepção e partiu-se para a segunda análise de trechos de protocolos verbais que pudessem revelar as concepções de análise dos indexadores, constatando-se a ocorrência de fusão de concepções e a existência de outras estratégias.

Outras estratégias observadas foram a associação com a linguagem do sistema, a influência da quantidade limite de termos e a influência da área especialista do indexador.

Observou-se que a associação com a linguagem, aparece nessa análise de forma mais específica, ou seja, sob a ótica de um processo de interpretação pelo resgate do CONCEITO na memória do indexador, tentando obter uma compreensão. A *associação com linguagem*, então, é característica da concepção orientada para o conteúdo, porque é demonstração de conhecimento prévio e, portanto, parte do processo de compreensão.

Com relação à fusão de concepções, observou-se que nenhum sujeito realizou apenas concepção simplista ou apenas orientada para a demanda, porém, os sujeitos 1 e 2 do Modelo de Leitura realizaram apenas a concepção orientada para o conteúdo. De acordo com a literatura, o uso de estratégias durante a leitura tende a um equilíbrio entre as estratégias perceptuais (características da concepção simplista) e conceituais (caracterizada pelo domínio da área de assunto em que a análise vai além da estrutura textual e superficial do documento). Todos os sujeitos, especialistas e não-especialistas no assunto, fazem uso de estratégias perceptuais e conceituais, portanto, realizam leitura compreensiva. Os especialistas, assim como os bibliotecários, também realizam a análise de acordo com a concepção simplista, sendo esta observada em todos os sujeitos, sempre combinada com a concepção orientada para o conteúdo.

Recomenda-se, desta maneira, que a leitura documentária em análise de assunto envolva tanto concepções como estratégias de leitura, a partir da perspectiva das áreas de interface: Lingüística, Lógica e Psicologia cognitiva. É necessário que a identificação de conceitos para a busca da tematicidade seja realizada a partir de análise conceitual; que a concepção simplista ou perceptual seja entendida como fase necessária para o leitor indexador atingir a compreensão; que a seleção de conceitos, durante a leitura para análise de assunto, seja orientada para a demanda sobre a perspectiva do usuário e da política de indexação do sistema de informação; que várias experimentações de leitura documentária em área especializadas sejam realizadas para conhecimento das tipologias docu-

mentárias e suas estruturas textuais, idealizando modelos de leitura para fins documentários.

2.3 O Contexto

Para os propósitos do estudo em questão e tendo em vista a abordagem interacionista assumida para a leitura documentária, o contexto, como anteriormente enunciado por Giasson (1993, p.40), pode ser distinguido em contexto físico, psicológico e social. No contexto físico encontram-se todas condições materiais para o desenvolvimento da leitura; no psicológico está o interesse e a intenção para a leitura e, no social, entendido como contexto sociocognitivo do indexador, está o conhecimento da situação comunicativa e de suas regras, implícitos no contexto do trabalho desenvolvido por indexadores em sistemas de informação.

2.3.1 Contexto psicológico: o objetivo da leitura documentária

O indexador tem como objetivo principal, representar o assunto de um documento por termos significativos para que seja possível sua recuperação pelos usuários interessados.

Examinando os objetivos específicos da leitura documentária, podemos inferir que são relacionados aos objetivos do sistema de informação e às necessidades dos usuários, consistindo basicamente em:

- Determinar o conteúdo principal do documento;
- Identificar e selecionar os conceitos para representar o conteúdo dos documentos.

Os objetivos são relacionados ao trabalho a ser desenvolvido pelo indexador e são pertinentes aos objetivos do sistema de informação. Dessa forma, o leitor passa a ser considerado um leitor profissional, porque os objetivos profissionais se sobrepõem aos objetivos pessoais. No caso da leitura documentária, o propósito consiste em extrair a informação relevante do texto, tendo em vista a sua posterior recuperação por um leitor interessado.

Como o indexador realiza uma leitura com objetivos profissionais, sua leitura documentária sofre a pressão da falta de tempo devido à grande quantidade de material que necessita ler para indexar. Sendo assim, o leitor-indexador utiliza, na leitura para fins de indexação, estratégias metacognitivas próprias de leitura documentária que lhe permitam atingir o seu objetivo. Para atingir a compreensão da leitura, o indexador utiliza-se de diversos processos existentes para a prática da leitura e, para a conclusão destes processos, apóia-se em estratégias visando a alcançar os objetivos.

Dentre os aspectos cognitivos envolvidos no processo de compreensão da leitura, tais como interesse, tarefa, objetivo, conhecimento, normas, opiniões ou atitudes; Van Dijk (1979), citado por Beghtol (1986), postula que o objetivo no processo de leitura representa o mais forte argumento na compreensão, pois, segundo o autor, o objetivo de leitura sobrepõe-se a qualquer tipo de estrutura textual. Supõe-se que a definição de objetivos para a leitura documentária atue como facilitador da compreensão e determinação do assunto do documento e deva fazer parte do programa de orientação, uma vez que, em protocolos verbais dos estudos de casos, a verbalização do objetivo durante o processo de leitura documentária foi raro. Compreender o texto pela leitura documentária, na análise de assunto, para melhor representar seu conteúdo e assim torná-lo disponível e conhecido aos usuários é um objetivo a ser melhor definido pelo indexador.

Em síntese, a leitura do indexador é guiada pelos seus objetivos e, dependendo de suas habilidades de leitor, conhecimentos prévios necessários à atividade de indexação, ele terá êxito.

2.3.2 Contexto físico

O contexto físico de sistemas de informação demonstram realidades diferenciadas na infra-estrutura física, importantes de serem consideradas no contexto do indexador. Torna-se necessária a análise dessas realidades na formação do indexador, tendo em vista que a forma de atuação profissional modifica-se de uma realidade para outra: em bibliotecas, o indexador realiza, além da indexação, atendimento ao usuário dentro de uma biblioteca aberta ao público; por outro lado, em serviços de informação, os indexadores realizam somente indexação, não atuam em um ambiente

de biblioteca e tampouco têm contato com o usuário. A influência dessa realidade física deverá refletir-se na política de indexação do sistema e na concepção de análise assumida pelo indexador.

2.3.3 Contexto sociocognitivo:

O *contexto sociocognitivo* aqui refere-se ao modelo mental do leitor indexador ao qual será necessário acrescentar os procedimentos de análise de assunto em indexação, a linguagem documentária do sistema e a política de indexação do sistema explicados por manual. O conjunto desses elementos deverá estar presente no conhecimento prévio do leitor profissional que realiza a leitura documentária.

a) o manual de indexação

O manual de indexação será considerado aqui como conteúdo a ser absorvido pelo indexador para a realização da leitura documentária dotada de estratégias específicas, tendo em vista seu contexto profissional. O conteúdo do Manual, então, expressa todo o contexto sociocognitivo do indexador constituído dos procedimentos de análise de assunto, linguagem documentária e política de indexação.

O manual de indexação, com objetivos, procedimentos, apresentação da linguagem e política de indexação, é parte necessária ao contexto sociocognitivo do indexador em leitura documentária, devendo contemplar os seguintes aspectos:

- Objetivos e propósito do sistema de informação;
- Apresentação dos procedimentos de indexação para identificação e seleção de conceitos articulados com o processo de leitura, incluindo um questionamento para esta finalidade, contendo exemplos em cada fase;
- Apresentação das linguagens documentárias alfabética e hierárquica adotadas para a representação dos conceitos selecionados, respectivamente, na indexação e classificação, esclarecendo aspectos de estrutura, vocabulário e configuração interna para uso, contendo exemplos;

- A especificação dos elementos constituintes da política de indexação do sistema: cobertura de assunto, critérios de seleção e aquisição dos documentos-fonte, nível de exaustividade e nível de especificidade.

b) a linguagem documentária

No âmbito da linguagem do sistema de informação, no que diz respeito à associação e validação, verificou-se com indexadores que a associação com linguagem permeia toda a operação e sempre está próxima (antes ou depois) da estratégia de identificação de conceitos. Esta associação com linguagem é um esquema do indexador não especialista acionado por seu conhecimento prévio, o que significa que o domínio do assunto é feito pela linguagem documentária do sistema;

Por outro lado, os indexadores que fazem associações com a linguagem documentária em poucos momentos, utilizando muito mais o próprio conhecimento prévio sobre o assunto, são especialistas e, portanto, dominam o assunto.

Considerando-se a linguagem documentária do sistema de informação como instrumento para domínio do assunto durante a compreensão de leitura, é importante haver: uma estrutura de conceitos visível e claramente compreensível ao indexador; explicações detalhadas e indispensáveis para o seu uso e aplicação; controle sobre as formas dos termos utilizados como descritores; diferenciadores tipográficos para diferenciação entre as relações de hierarquia, associações e equivalência.

Além disso, recomenda-se que o indexador deve ter conhecimento sobre a estrutura da linguagem e seu vocabulário para aumentar seu conhecimento prévio e, se possível, aumentar o contato com usuários em oportunidades variadas: eventos específicos, reuniões de colegiado, reuniões de grupos de pesquisa; acompanhar a elaboração de trabalhos acadêmicos, estratégias de busca, etc.

c) a política de indexação

Considerando que a política de indexação adotada influenciará diretamente a indexação que, por sua vez, será refletida na recuperação da informação, esta política do serviço de análise deve estar clara e objetivamente expressa em seu manual de indexação para que os indexadores tenham dela conhecimento e compreensão.

Os elementos de política de indexação que influenciam diretamente na leitura do indexador são a *exaustividade* e a *especificidade,* porque são essas variáveis que *interferem na escolha dos termos para indexação:*

- exaustividade: o indexador, no momento da leitura, deverá estar ciente do *número* de descritores que pode extrair de cada documento, obrigando-o a escolher entre um e outro, fazendo assim a seleção. Entretanto, é necessário que o sistema de informação estabeleça um número (mínimo e/ou máximo) de descritores a serem escolhidos. A ausência desta delimitação apresentará variabilidade no número de descritores escolhidos por cada indexador, gerando interconsistência.
- especificidade: caso seja recomendação do sistema de informação que o indexador seja o mais específico possível, será necessário que ele leia o documento, tendo em mente o nível de especificidade exigida pelo sistema.

Referências

ALBRETCHTSEN, H. Subject analysis and indexing: from automated indexing to domain analysis. *The Indexer*, London, v. 18, n. 4, p. 219-24, oct. 1993.

BEAUGRANDE, R. *New foundations for a science of text and discourse*: cognition, communication and freedom of access to knowledge and society. Norwood: Alex, 1997.

BEGHTOL, C. Bibliographic classification theory and text linguistics: aboutness analysis, interxtuality and the cognitive act of classifying documents. *J. Doc.*, London, v. 42, n. 2, p. 84-113, 1986.

BROWN, A. L. Metacognitive development and reading. In: SPIRO, et al. (Org.). *Theorical issues in reading comprehension*. New Jersey: Lauwrence Erlbaum Associates, 1980. p. 453-481.

CAVALCANTI, M. C. *I-n-t-e-r-aç-ã-o leitor-texto*: aspectos de interpretação pragmática. Campinas: UNICAMP, 1989. 271 p.

FUJITA, M. S. L. *A leitura em análise documentária.* 1998. 184 f. Relatório final (Projeto Integrado de Pesquisa) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista; CNPq, Marília.

______. *Leitura em análise documentária*. 1999c. 123 f. Relatório parcial (Projeto Integrado de Pesquisa) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista; CNPq, Marília.

______. *A leitura documentária do indexador*: aspectos cognitivos e lingüísticos influentes na formação do leitor profissional. 321 f. Tese (Livre-Docência em Análise Documentária e Linguagens Documentárias Alfabéticas) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2003.

GIASSON, J. *A compreensão na leitura*. Lisboa: Asa, 1993.

KLEIMAN, A. *Texto e leitor*: aspectos cognitivos da leitura. 7.ed. Campinas: Pontes, 2000.

______. *Leitura*: ensino e pesquisa. Campinas: Fontes, 1989.

KOBASHI, N. Y. *A elaboração de informações documentárias:* em busca de uma metodologia. 1994. 195 f. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo.

KOCH, I.G.V. *Desvendando os segredos do texto*. São Paulo: Cortez, 2002.

LEFFA, V. J. *Aspectos da leitura*. Porto Alegre: Sagra. 1996.

______. Análise de assunto: concepções. *Revista de Biblioteconomia de Brasília*, Brasília, v. 20, n. 2, p. 215-226, jul./dez, 1996.

PINTO, M.; GÁLVEZ, C. *Análisis documental de contenido:* processamiento de información. Madrid : Síntesis, [1996].

RUMELHART, D.E. Toward an interactive model of reading. In: DORMICI, S. (Org). *Attention and performance XL*. [s.l.]: Lawrence Esrlbaum Associates, 1977.

VAN DIJK, T. A *La ciência del texto:* um enfoque interdisciplinário. Tradução de Sibila Hunzinger. Barcelona: Paidos, 1992. Tradução de: Tekstwetenschap. Een Interdisplinaire inleiding

Sobre os autores

Ana Lúcia Terra

Professora Adjunta no Instituto Politécnico do Porto
anaterra@iscap.ipp.pt

É doutorada em Ciências Documentais pela Universidade de Coimbra, Licenciada em História e Mestre em História Moderna pela mesma instituição. Pós-graduada em Ciências Documentais (Opção Arquivo e Biblioteca/Documentação) pela Universidade de Coimbra. É Professora Adjunta no Instituto Superior de Contabilidade e Administração do Porto. Desde 2002, leciona na área da Ciência da Informação, na Licenciatura em Ciências e Tecnologias da Documentação e Informação, de que foi coordenadora no período de 2009-2016. Também leciona no Mestrado em Informação Empresarial, de que é diretora desde a sua primeira edição. Tem colaborado igualmente na Pós-Graduação em Gestão de Bibliotecas Escolares. É investigadora integrada do CIC.Digital, no pólo da Universidade do Porto. Foi galardoada com o Prémio Raúl Proença, da Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas, em 2009. É sócia de diversas associações profissionais e científicas: Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas, International Council on Archives, International-Society for Knowledge Organization (ISKO) – Capítulo Ibérico e da SOPCOM – Associação Portuguesa de Ciências da Comunicação. É coordenadora do Grupo de Trabalho de Ciência da Informação da SOPCOM. As suas áreas de investigação abrangem a organização e representação da informação, o comportamento informacional e a gestão da informação.

Brisa Pozzi de Sousa

Docente do curso de Biblioteconomia - UNIRIO
brisapozzi@gmail.com

Bacharel em Biblioteconomia pela Universidade Federal do Espírito Santo (UFES) e mestre em Ciência da Informação pela Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESPcâmpus de Marília). Atualmente é doutoranda no Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Minas Gerais (UFMG) e docente nos cursos de bacharelado e licenciatura em Bi-

blioteconomia da Universidade Federal do Estado do Rio de Janeiro (UNIRIO). Integra os grupos de pesquisa Organização do Conhecimento para Recuperação da Informação, liderado pelo professor Marcos Luiz Cavalcanti de Miranda (UNIRIO) e Fundamentos Teóricos, Metodológicos e Históricos da Organização da Informação, liderado pela professora Cristina Dotta Ortega (UFMG).Seu foco centrana Biblioteconomia e Ciência da Informação, nas temáticas que envolvem o contexto da Organização do Conhecimento, destacando-se: análise de assunto, indexação, classificação, modelagem conceitual e catalogação.

Carlos Cândido de Almeida

Pesquisador nível 2 do CNPq
Departamento de Ciência da Informação
Faculdade de Filosofia e Ciências
UNESP – Campus de Marília
carlosalmeida@marilia.unesp.br

Docente do Departamento de Ciência da Informação da Universidade Estadual Paulista, campus de Marília, com atuação nos cursos de Graduação em Arquivologia e Biblioteconomia e no Programa de Pós-Graduação em Ciência da Informação - UNESP. Pós-doutor em Biblioteconomía y Documentación pela Universidad de Zaragoza, Espanha, doutor em Ciência da Informação pela Universidade Estadual Paulista, mestre em Ciência da Informação pela Universidade Federal de Santa Catarina e bacharel em Biblioteconomia pela Universidade Estadual de Londrina. Líder do grupo de pesquisa Fundamentos Teóricos da Informação. Bolsista de produtividade em pesquisa CNPq nível 2. É membro das sociedades científicas: Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação ANCIB e do Capítulo Brasileiro da International Society for Knowledge Organization ISKO, no Brasil. É parecerista de agências de fomento e membro de comitês científicos de eventos e periódicos em Ciência da Informação, no Brasil e no exterior. Realiza pesquisas nas áreas: epistemologia da ciência da informação, semiótica, comunicação, organização da informação, organização do conhecimento e mediação.

Daniela Majorie Akama dos Reis

Doutoranda em Ciência da Informação
UNESP – Campus de Marília
danielamajorie@yahoo.com.br

Bacharel em Biblioteconomia pela Universidade Estadual Paulista "Júlio de Mesquita Filho" – UNESP com bolsa de iniciação científica CNPq/PIBIC. Mestra em Ciência da Informação pela Universidade Estadual Paulista "Júlio de Mesqui-

ta Filho" – UNESP com bolsa FAPESP, com dissertação intitulada: A importância da observação da estrutura textual durante a catalogação de assunto de livros científicos em bibliotecas universitárias: uma análise realizada a partir da técnica de Protocolo Verbal. Doutoranda em Ciência da Informação pela Universidade Estadual Paulista "Júlio de Mesquita Filho" – UNESP, com previsão para conclusão e defesa da Tese em 2019. Estuda atualmente aspectos da Semiótica em tratamento de livros na Organização do Conhecimento.

Dulce Amélia de Brito Neves

Professora Titular do Departamento de Ciência da Informação da UFPB.
damelia1@gmail.com

Doutora em Ciência da Informação pela UFMG.Docente permanente do Programa de Pós-graduação em Ciência da Informação da UFPB, na linha Organização, Acesso e Uso da Informação. Coordena os Grupos de Pesquisa: Leitura, organização, representação, produção e uso da informação e Representação temática da informação em Arquivística. Membro da: Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação ANCIB e do Capítulo Brasileiro da InternationalSociety for Knowledge Organization - ISKO, no Brasil e no exterior da InternationalSociety for KnowledgeOrganization - ISKO. Faz parte do Corpo Editorial da Revista Informação & Sociedade: Estudos e Revista Biblionline, da Revista Ciência da Informação, Revista Encontros Bibli (UFSC), Informação & Informação, Revista Brasileira de Pós-graduação (Capes), JournalofInformation Science, Revista ACB (Associação Catarinense de Bibliotecários), Ponto de Acesso. Consultora Had Hoc da Capes. Áreas de interesse Ciência da Informação, com ênfase em Biblioteconomia e Arquivologia, atuando principalmente em: Representação da Informação, Organização da Informação. Indexação, Metacognição, Leitura em Unidades de Informação, Comportamento Informacional. Membro da ISKO Brasil

Érica Fernanda Vitorini

Doutoranda em Ciência da Informação
UNESP – Campus de Marília
erica.vitorini@marilia.unesp.br

Doutoranda em Ciência da Informação no PPGCI na Universidade Estadual Paulista Julio de Mesquita Filho - Campus de Marília. Membro do Grupo de pesquisa Análise Documentária. Mestra em Ciência da Informação no PPGCI na Universidade Estadual Paulista Julio de Mesquita Filho - Campus de Marília em 2015, com o projeto Uso da linguagem documentária na busca da informação em bibliotecas universitárias: a perspectiva dos deficientes visuais. Bacharel de

Biblioteconomia e Ciência da Informação na Universidade Federal de São Carlos-Campus de São Carlos em 2010. Bolsista FAPESP em 2010 com o projeto de Iniciação Científica A linguagem documentária vista pelo conteúdo em catálogos coletivos de bibliotecas universitárias: avaliação qualitativa-sociocognitiva pela perspectiva do bibliotecário indexador.

José Carlos Francisco dos Santos

Discente do Programa de Pós-Graduação em Ciência da Informação
UNESP – Campus de Marília
Docente das Faculdades Integradas do Vale do Ivaí - Ivaiporã - PR
jose.santos@marilia.unesp.br

Doutorando em Ciência da Informação na Universidade Estadual Paulista "Júlio de Mesquita Filho" – Unesp, Mestre em Ciência da Informação pela Universidade Estadual de Londrina - UEL (2015). Docente e Coordenador de Estágio Supervisionado do Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas das Faculdades Integradas do Vale do Ivaí. Gerente de Projetos do Esap - Instituto de Estudos Avançados e Pós-Graduação. Como pesquisador, atua nos Grupos de Pesquisa "Análise Documentária" (na categoria Estudante) da Unesp, "Controle de Vocabulário em periódicos científicos eletrônicos: proposta de implantação da VCPC Tools no periódico "Discursos Fotográficos" (como Colaborador *ad hoc*). Atua na área de Ciência da Informação, Organização do Conhecimento, com ênfase em Controle de Vocabulário em periódicos científicos eletrônicos.

Juan-Francisco Torregrosa Carmona

Profesor Titular de Documentación Informativa
Universidad Rey Juan Carlos (Madrid, España)
juanfrancisco.torregrosa@urjc.es

Docente e investigador de la Facultad de Ciencias de la Comunicación de la Universidad Rey Juan Carlos (URJC, Madrid), en la que trabaja desde 2004 a tiempo completo, en grado y posgrado, presencial y semipresencial. Es Coordinador Académico (Vicedecano) del Grado de Comunicación Audiovisual y miembro de la Cátedra UNESCO de Investigación en Comunicación de la URJC. Doctor por la Complutense, UCM (en Ciencias de la Información, sobresaliente *cum laude* por unanimidad), donde fue personal docente e investigador. Diploma de Estudios Avanzados en Estructura, Tecnología y Tratamiento de la Información. Diploma de Aptitud Pedagógica en Lengua Española y Literatura.Licenciado en Periodismo y Licenciado en Derecho, su formación principal ha correspondido a la UCM y a la UNED (Universidad Nacional de Educación a distancia). Du-

rante años trabajó como redactor, columnista y documentalista en prensa, radio y televisión, así como en la Dirección de Comunicación de empresas privadas y de la Administración Pública. Perito judicial para cadenas nacionales de televisión públicas y privadas, ha impartido clase en programas de doctorado y títulos de máster de instituciones y medios de comunicación y en universidades tanto españolas como internacionales. Dentro de estas últimas ha sido conferenciante e investigador en países de Europa, América Latina (Universidad Austral de Chile e Instituto Tecnológico de Monterrey, México) y Asia-Pacífico (universidades de Pekín y *Waseda Universtiy*, de Tokio).

Luciana Beatriz Piovezan
Doutoranda em Ciência da Informação
Programa de Pós-Graduação em Ciência da Informação
Universidade Estadual Paulista – UNESP – Campus de Marília
lbpiovezan@gmail.com
Doutoranda em Ciência da Informação, vinculada à linha de pesquisa Produção e Organização da Informação do PPGCI da Universidade Estadual Paulista – UNESP, com o tema Linguagem de indexação interoperável para bibliotecas universitárias. Possui mestrado em Ciência da Informação (2015) e graduação em Biblioteconomia (2011) pela UNESP. Foi membro da Sub-Comissão da CADA para elaboração de Vocabulário Controlado de termos arquivísticos da UNESP - VTArq (2013-2014). É membro do Grupo de Política de Indexação e do Grupo de Linguagem da Rede de bibliotecas da UNESP. Desenvolve pesquisas na área de Ciência da Informação, tendo como principais interesses as temáticas indexação, política de indexação, linguagem de indexação e leitura documental para indexação.

María del Carmen Agustín Lacruz
Profesora Titular
Universidad de Zaragoza (España)
cagustin@unizar.es
Profesora e investigadora en la Universidad pública de Zaragoza. Doctora en Filosofía y Letras(sobresaliente cum laude), licenciada en Filología Hispánica y Diplomada en Biblioteconomía y Documentación,trabajó también durante ocho años como bibliotecaria en la Red de Bibliotecas Públicas municipales de Zaragoza. Actualmente es Directora del Departamento de Ciencias de la Documentación e Historia de la Ciencia y coordinadora del Programa de doctorado interuniversitario (Universidad de Barcelona-Universidad de Zaragoza) en Información y Documentación. Imparte docencia en los Grados de Información y Documen-

tación, y Periodismo, así como en el master en Consultoría en Información y Comunicación Digital. Es autora de un centenar de trabajos sobre bibliotecas, lectura, inclusión digital, análisis de imágenes, fuentes de información y organización del conocimiento. Ha impartido cursos y seminarios en diferentes universidades españolas, portuguesas (Évora, Oporto, Instituto Politécnico de Oporto) y brasileñas (UNESP, UFSC y UFSCar).

Mariângela Spotti Lopes Fujita

Bolsista PQ-1C/CNPq
Docente Permanente Programa de Pós-Graduação em Ciência da Informação
UNESP – Campus de Marília
fujita@marilia.unesp.br

Doutora em Ciências da Comunicação pela Universidade de São Paulo (1992), Livre Docente (2003) em Análise Documentária e Linguagens Documentárias Alfabéticas, Professora Titular da Universidade Estadual Paulista "Júlio de Mesquita Filho" – UNESP no período de 2010 a 2017. É docente permanente na Pós-Graduação na linha de pesquisa "Produção e Organização da Informação" do Programa em Ciência da Informação da UNESP. Como Pesquisadora atua nos Grupos de Pesquisa "Representação Temática da Informação" (líder), "Leitura, organização, representação, produção e uso da informação – UFPB" e "Organização do conhecimento para disseminação da informação-UFSCar". Desenvolve atividades de pesquisa na UNESP com bolsa de Produtividade em Pesquisa do CNPq nível 1C. É membro das Sociedades Científicas de sua especialidade: Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação ANCIB e do Capítulo Brasileiro da International Society for Knowledge Organization ISKO, no Brasil e no exterior da International Society for Knowledge Organization ISKO. Atualmente é Assessora e Presidente do Comitê de área de Comunicação, Artes e Ciência da Informação do CNPq, Supervisora do Instituto de Políticas Públicas de Marília e Presidente da Comissão Permanente de Publicações da Faculdade de Filosofia e Ciências da UNESP. É parecerista ad hoc de agências de fomento e participa como revisora e membro de Comitês Científicos de eventos e periódicos científicos em Ciência da Informação no Brasil e no exterior. Atua na área de Ciência da Informação, com ênfase em Indexação, Leitura Documentária para indexação, utilizando a metodologia introspectiva de Protocolo Verbal em diferentes modalidades. Realiza pesquisas sobre os tema de política de indexação e Linguagem de indexação para bibliotecas.

Milena Polsinelli Rubi
Bibliotecária
Universidade Federal de São Carlos (UFSCar) – Campus Sorocaba
milenarubi@ufscar.br

Doutora em Ciência da Informação pela Faculdade de Filosofia e Ciências, Universidade Estadual Paulista (UNESP), Marília, SP (2008). Possui graduação em Biblioteconomia pela UNESP (2000) e mestrado em Ciência da Informação também pela UNESP (2004). Atuou como professora substituta do Departamento de Ciência da Informação da Universidade Federal de São Carlos (UFSCar). Foi bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior durante o mestrado e doutorado e bolsista de iniciação científica do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Tem experiência na área de Ciência da Informação, atuando principalmente nos seguintes temas: indexação, política de indexação, análise documentária, protocolo verbal e estratégias de leitura. Atualmente, é bibliotecária na Universidade Federal de São Carlos (UFSCar), Campus Sorocaba.

Roberta Caroline Vesu Alves
Mestre e Doutora em Ciência da Informação - Unesp de Marília
UNESP – Campus de Marília
robertavesu@gmail.com

Mestre em Ciência da Informação (2008) pela Universidade Estadual Paulista "Júlio de Mesquita Filho" – UNESP, Bibliotecária Escolar do Centro Educacional do Serviço Social da Indústria – SESI de Guararapes-SP (2009-2010), Bibliotecária Jurídica (2010-2012) e Doutora em Ciência da Informação (2016) pela UNESP com bolsa CAPES e, posteriormente, FAPESP. Bolsista CNPq-AT (2007) do projeto "Abordagem cognitiva da leitura documentária no tratamento documentário com uso do Protocolo Verbal", com orientação da Profª. Drª. Mariângela Spotti Lopes Fujita. É membro da Sociedade Científica de sua especialidade: Capítulo Brasileiro da International Society for Knowledge Organization – ISKO, no Brasil. Atuou como Professora Substituta do Departamento de Ciência da Informação da UNESP em 2016 e 2017. Atua na área de Ciência da Informação, com ênfase em Análise Documental e indexação com leitura e identificação de temas para representação de conteúdo dos textos literários infanto-juvenis, Análise Documental para representação e elaboração de resumos de textos narrativos ficcionais infanto-juvenis, considerando aspectos de *aboutness*, Semiótica de Greimas, Análise do Discurso, Percurso Gerativo de Sentido, teorias da Narrativa, Linguística Textual e elementos textuais e cognitivos da Leitura Documentária para indexação e representação, além da Análise Documental no âmbito do Tratamento Temático da Informação, Ciência da Informação e Organização do Conhecimento.

Walter Moreira
Departamento de Ciência da Informação
Faculdade de Filosofia e Ciências
UNESP – Campus de Marília
walter.moreira@marilia.unesp.br

Doutor em Ciência da Informação pela Universidade de São Paulo (2010), Professor Assistente Doutor da Universidade Estadual Paulista "Júlio de Mesquita Filho". Atua nos cursos de graduação em Arquivologia e Biblioteconomia e no Programa de Pós-graduação em Ciência da Informação, na linha de pesquisa "Produção e Organização da Informação". Como pesquisador atua nos grupos de pesquisa "Análise Documentária - UNESP" (líder) "Representação Temática da Informação - UNESP", "Linguagem, discurso e organização do conhecimento – UNESP". Desenvolve pesquisas sobre sistemas de organização do conhecimento, linguagens documentárias, teoria da classificação, terminologia e ontologias.

Wendia Oliveira de Andrade
Doutoranda em Ciência da Informação
Mestra em Ciência da Informação
UFPB – Campus de João Pessoa
wendia2810@gmail.com

Doutoranda (2015-2019) e Mestra distinta (2014) em Ciência da Informação pela Universidade Federal da Paraíba, com a temática de usuários de arquivo. É bibliotecária (2010) e atuou como Professora substituta na Universidade Estadual da Paraíba (2014-2016), ministrando disciplinas no curso de Bacharelado em Arquivologia, a saber: Representação da Informação Arquivística; Arquivos Permanentes; Reprodução de documentos; Gestão da Informação Arquivística e Metodologia da Pesquisa (ênfase na construção de trabalhos acadêmicos). Regularmente oferece oficinas e palestras, com temática sobre as normas vigentes da Associação Brasileira de Normas Técnicas (ABNT) para elaboração de trabalhos acadêmicos e também sobre Tesauros, como instrumentos de representação e recuperação de informação em arquivos e bibliotecas. Tem experiência na área de Arquivologia, com ênfase na organização e recuperação da informação (SRI) de arquivos jurídicos. Tem interesse em temáticas como usuários de arquivo, acessibilidade e uso de SRI, tanto em arquivos quanto em bibliotecas; disseminação e recuperação da informação; linguagens documentárias; representação temática da informação arquivística: tesauros e unitermos. Na área da Ciência da Informação, tem se dedicado aos estudos sobre o conceito de informação, relacionando-o como a Arquivologia.

No contexto da indexação de assunto, o tema leitura documentária, investigado nesta obra, direciona-se aos desafios e perspectivas científicas contemporâneos da organização da informação. Reúne pesquisadores preocupados com os desdobramentos e implicações do processo de leitura documentária em Ciência da Informação e que têm feito desta preocupação um elemento da sua reflexão. Os caminhos percorridos no entorno da leitura documentária viabilizam a sua compreensão e favorecem a transposição dos limites entre a teoria e a prática complexa da indexação, com fins de representação e recuperação documentárias. Essa compreensão rompe com o processo de obviedade na análise e representação de assunto, conferindo ao estudo da leitura em indexação um lugar de destaque.

Paula Regina Dal´Evedove

ISBN 978-85-7983-916-0